

STIC-Biotech/ChemLib

From: Ramirez, Delia
Sent: 78901 Friday, October 06, 2006 4:26 PM
To: STIC-Biotech/ChemLib
Subject: 10/824581

Hi,

I would like to request the following search: SEQ ID NO:4 in the protein databases (commercial & interference).

Please provide a printout of the results.

Thank you very much,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

4- 161 AA

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

Date completed: _____

Searcher: Beverly e 2528

Terminal time: _____

Elapsed time: _____

CPU time: _____

Total time: _____

Number of Searches: _____

Number of Databases: _____

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

Vendors

_____ IG

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

_____ Other CBN

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM protein - protein search, using sw model
 Run on: October 12, 2006, 05:31:52 ; Search time 199 Seconds
 (without alignments)
 369.909 Million cell updates/sec
 Title: US-10-824-581-4
 Perfect score: 882
 Sequence: 1 MENFIGSHMIYENGWEYE.....EVISKAPYEGMTDIRAGEL 161
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 2589679 seqs, 457216429 residues
 Total number of hits satisfying chosen parameters: 2589679
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : A_Geneseq 8:
 1: Geneseqp1980s:
 2: Geneseqp1980s:
 3: Geneseqp2000s:
 4: Geneseqp2001s:
 5: Geneseqp2002s:
 6: Geneseqp2003as:
 7: Geneseqp2003bs:
 8: Geneseqp2004s:
 9: Geneseqp2005s:
 10: Geneseqp2006s:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	882	100.0	161	ADG28813	Bacillus
2	882	100.0	161	ADU03637	Bacillus
3	777	88.1	161	AY28582	Synthetic
4	697	79.0	176	AB55283	Lactococ
5	675	76.5	183	ADG96607	E. faeciu
6	666.5	75.6	174	ADG28811	L. plantar
7	666.5	75.6	174	ADU03635	Lactobaci
8	499	56.6	168	AB064330	Klebsiell
9	91.5	10.4	1574	ADN17494	Bacterial
10	88.5	10.0	382	ADA36451	Acinetoba
11	87	9.9	352	ABU16806	Protein e
12	83.5	9.5	497	AB94833	Human pro
13	83.5	9.5	497	ABP69347	Human pol
14	83.5	9.5	508	ABP69348	Human pol
15	83	9.4	652	ADT58254	Plant pol
16	83	9.4	962	AEA28265	Streptoco
17	81	9.2	233	ABY49036	Listeria
18	81	9.2	740	AAV99411	Human PRO
19	81	9.2	740	AA866160	Protein o
20	80.5	9.1	410	ABP77891	N. gonorr
21	80	9.1	800	ABU02167	S. pneumo
22	80	9.1	800	ABP81446	Streptoco
23	80	9.1	946	ADT50135	S_pneumon

24	80	9.1	1035	9	AEA28264	Streptoco
25	80	9.1	1368	8	ADS22521	Bacterial
26	79	9.0	217	3	AAG61391	Arabidops
27	79	9.0	217	3	AAG11399	Arabidops
28	79	9.0	230	3	AAG11398	Arabidops
29	79	9.0	230	3	AAG61390	Arabidops
30	79	9.0	251	3	AAG11397	Arabidops
31	79	9.0	251	3	AAG61389	Arabidops
32	79	9.0	251	8	ADT56317	Plant pol
33	79	9.0	354	6	ABU37073	Protein e
34	79	9.0	740	4	ABG11840	Novel hum
35	78.5	8.9	731	5	ABP65510	Bifidobac
36	78.5	8.9	1251	5	ABG70358	Novel hum
37	78.5	8.9	1602	5	ABG70359	Novel hum
38	78.5	8.9	1676	2	AAK77604	Pro-C5 po
39	78.5	8.9	1676	9	AD280439	CSA anaph
40	78.5	8.9	1676	9	AEC74697	Synthetic
41	78.5	8.9	1676	9	AED60676	Human com
42	78.5	8.9	1680	8	ADP79661	Mouse com
43	78.5	8.9	1680	9	AED60677	Mouse com
44	78	8.8	452	4	AB558705	Drosophil
45	78	8.8	3135	2	AAK57474	P. faicip

ALIGNMENTS

RESULT 1
 ADG28813
 ID ADG28813 standard; protein; 161 AA.
 AC ADG28813;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Bacillus subtilis para-hydroxycinnamic acid decarboxylase (pdc2) protein.
 XX
 KW para-hydroxystyrene; pHS; tyrosine ammonia lyase;
 KW para-hydroxycinnamic acid decarboxylase; resin; permselective membrane;
 KW TAL; PDC; polymer support; coating; ink; pdc2; enzyme.
 XX
 OS Bacillus subtilis.
 XX
 PN WO2003099233-A2.
 XX
 PD 04-DEC-2003.
 XX
 PF 20-MAY-2003; 2003WO-US017926.
 XX
 PR 23-MAY-2002; 2002US-0383450P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PI Ben-Bassat A, Qi WW, Sartaslani FS, Tang X, Vannelli T;
 DR WPI: 2004-071031/07.
 DR N-PSDB; ADG28812.
 PT Production of para-hydroxystyrene comprises contacting a recombinant host
 PT cell comprising a gene encoding tyrosine lyase and a gene encoding
 PT parahydroxycinnamic acid decarboxylase, with a fermentable carbon
 PT substrate.
 XX Claim 18; SEQ ID NO 6; 81pp; English.
 PS The invention relates to a novel method for production of para-
 CC hydroxystyrene (pHS) comprising contacting a recombinant host cell with a
 CC fermentable carbon substrate, where the recombinant host comprises at
 CC least one gene encoding a polypeptide having tyrosine ammonia lyase (TAL)
 CC activity and at least one gene encoding a polypeptide having para-
 CC hydroxycinnamic acid decarboxylase (PDC) activity and growing the
 CC recombinant cell to produce pHS. The method of the invention may be
 CC useful for the production of pHS, having applications in the manufacture

CC of resins, permselective membranes, polymer supports, coatings and inks.
 CC The method uses inexpensive materials such as carbohydrates or sugars and
 CC is more efficient, relative to existing methods, in the production of pHs
 CC with limited by-products. The current sequence is that of the Bacillus
 CC subtilis para-hydroxycinnamic acid decarboxylase (pdc2) protein of the
 CC invention.

XX Sequence 161 AA;

Query Match 100.0%; Score 882; DB 8; Length 161;
 Best Local Similarity 100.0%; Pred. No. 2.5e-86;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MENFGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMWAGRWVRDQEVNIVKLTGVYKV 60
 DB 1 MENFGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMWAGRWVRDQEVNIVKLTGVYKV 60

OY 61 SWTEPTGTVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLKESREKYETYP 120
 DB 61 SWTEPTGTVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLKESREKYETYP 120

OY 121 KYVPEFAEITFLKNEGVNDNEEVISKAPYEGMTDDIRAGRL 161
 DB 121 KYVPEFAEITFLKNEGVNDNEEVISKAPYEGMTDDIRAGRL 161

RESULT 2.

ADU03637
 ID ADU03637 standard; protein; 161 AA.

AC ADU03637;

XX 13-JAN-2005 (first entry)

DE Bacillus subtilis phenolic acid decarboxylase.

XX para-hydroxystyrene; para-hydroxycinnamic acid decarboxylase;

KW para-acetoxystyrene; resin; elastomer; adhesive; coating;

KW automotive finish; ink; electronic material;

KW para-coumaric acid decarboxylase; pdc2; enzyme.

XX Bacillus subtilis.

OS WO2004092344-A2.

PN 28-OCT-2004.

XX 14-APR-2004; 2004WO-US011510.

XX 14-APR-2003; 2003US-0462827P.

PR 17-FEB-2004; 2004US-0777446.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Haynie SL, Ben-Bassat A, Lowe DJ, Huang LL;

XX WPI; 2004-804446/79.

DR N-PSDB; ADU03636.

XX Producing para-hydroxystyrene as monomer for production of e.g. resins,

PT comprises providing enzyme source having para-hydroxycinnamic acid

PT decarboxylase activity.

XX Claim 1; SEQ ID NO 4; 57pp; English.

XX The invention relates to a method of producing para-hydroxystyrene by:

CC (a) providing an enzyme source having para-hydroxycinnamic acid

CC decarboxylase activity; (b) contacting the enzyme source with para-

CC hydroxycinnamic acid in a biphasic reaction medium comprising an aqueous

CC phase and an extractant, which is a water-immiscible organic solvent

CC consisting of toluene, methyl decanoate, 2-undecanone, dichloromethane,

CC hexane, 2-decanol, 4-decanol, 3-decanone, 4-decanone, 1-nonanol, 2-

CC nonanol, and/or 2-heptanol, to form para-hydroxystyrene, which is

CC extracted into the extractant of the biphasic reaction medium; (c)
 CC separating the extractant from the aqueous phase; and (d) optionally
 CC recovering the para-hydroxystyrene from the extractant. The method is
 CC useful for the production of para-hydroxystyrene or its derivatized
 CC compound, which is para-acetoxystyrene, useful as monomer for the
 CC production of resins, elastomers, adhesives, coatings, automotive
 CC finishes, inks and electronic materials, and additives in elastomer and
 CC resin formulations. High product yields of para-hydroxystyrene are
 CC obtained due to the decreased exposure of the enzyme source to the
 CC inhibitory product, which is extracted into the organic phase of the
 CC biphasic reaction medium. It provides excellent preservation of enzymatic
 CC activity and reuse of the enzyme source for many reaction cycles. This
 CC sequence corresponds to a Bacillus subtilis phenolic acid decarboxylase
 CC (pdc2) protein used in the method of the invention.

XX Sequence 161 AA;

Query Match 100.0%; Score 882; DB 8; Length 161;
 Best Local Similarity 100.0%; Pred. No. 2.5e-86;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MENFGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMWAGRWVRDQEVNIVKLTGVYKV 60
 DB 1 MENFGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMWAGRWVRDQEVNIVKLTGVYKV 60

OY 61 SWTEPTGTVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLKESREKYETYP 120
 DB 61 SWTEPTGTVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLKESREKYETYP 120

OY 121 KYVPEFAEITFLKNEGVNDNEEVISKAPYEGMTDDIRAGRL 161
 DB 121 KYVPEFAEITFLKNEGVNDNEEVISKAPYEGMTDDIRAGRL 161

RESULT 3

AA28582

ID AA28582 standard; protein; 161 AA.

XX AA28582;

XX 09-NOV-1999 (first entry)

DE Synthetic ferulic acid decarboxylase clone pGS97b1.

XX Ferulic acid decarboxylase; secondary metabolic pathway; sinapine; taste;

KW seed; phenylpropanoid pathway; insect attraction; disease tolerance;

KW genetically transformed plant; betaine.

XX Synthetic.

OS Bacillus pumilus.

XX WO9937786-A2.

XX 29-JUL-1999.

XX 22-JAN-1999; 99WO-CA000056.

XX 22-JAN-1998; 98US-0072156P.

PR 23-JAN-1998; 98US-00012453.

XX (CANA) NAT RES COUNCIL CANADA.

XX Georges F, Dong J, Keller WA, Hussain AAK, Selvaraj G, Datla RSS;

XX WPI; 1999-493953/41.

DR N-PSDB; AA210824.

XX Genetic modification of levels of secondary metabolic compounds in plants

PT useful for altering, e.g. phenolic content and sugar content in seeds.

XX Example 7; Fig 18; 145pp; English.

XX This sequence is a synthetic version of the ferulic acid decarboxylase

CC The present invention is related to a Lactococcus lactis nucleotide

The invention relates to a method of producing para-hydroxystyrene by:

XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 1574 AA;

Query Match 10.4%; Score 91.5; DB 8; Length 1574;
 Best Local Similarity 22.0%; Pred. No. 5.2;
 Matches 39; Conservative 32; Mismatches 69; Indels 37; Gaps 6;

QY 4 FIGSHMTYVNGWEYIYKNDHTIDYRHSWAGRWQDQVNVKLTGKYSWT 63
 Db 134 YFESYLVIIEPNEEBEKKFDEHTIP--LNDG-ISTKWKVLHYVNEEFEE----- 182
 QY 64 EPTGTVSLNFMPEKRMHGI-----IFFPKWVHEHPITVCYQNDHIDLMKES 112
 Db 183 -----KYAFTTDEKYEHGNGAEILKVLKLDLDAYSRLKLEIVKPYSIGEDLGEI 235
 QY 113 REK-----YETPKYVWPEP-----AEITFLKNEGVNDEEISKAPYEGMTDDIRAGRL 161
 Db 236 EQKYNLYQXLIKVIADDFRAYGVEIKGLEDHGLSLQAIHRIILNBSLYLNVERGEI 292

RESULT 10

ADA36451
 ID ADA36451 standard; protein; 382 AA.

AC ADA36451;

DT 20-NOV-2003 (first entry)

XX Acinetobacter baumannii protein #3612.

XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.

XX Acinetobacter baumannii.

XX US6562958-B1.

XX 13-MAY-2003.

XX 04-JUN-1999; 99US-00328352.

XX 09-JUN-1998; 98US-0088701P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton G, Bush D;

XX WPI; 2003-576092/54.

DR N-PSDB; ADA32325.

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.

XX Example; SEQ ID NO 7738; 328pp; English.

XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.

XX SQ Sequence 382 AA;

Query Match 10.0%; Score 88.5; DB 6; Length 382;
 Best Local Similarity 24.8%; Pred. No. 1.6;
 Matches 39; Conservative 23; Mismatches 56; Indels 39; Gaps 7;

QY 13 YENGWEYIYKNDHTIDYRHSWAGRW-----VRDQEVNIVKLTGCV--YKVSWTEP 65

Db 10 YFNGRRNSLSLKQHPDIREII---MKAAPYDRGDIRIEDIPPEVTPGTGIVKAVCGI 66

QY 66 TGTDSVSLNFMPEKRMHGIFFPKWVHEHPITVCYQNDHIDLMKESREKYETPKYVVP 125

Db 67 CGTDL-----HEFMEGPPIFPCCGHPHP-----ISGESAPITMGH 101

QY 126 EFAEITFLKNEGVNDEEV-----ISKAPYEGMTDDIRAG 159

Db 102 EFGSVYAVGEGVDYIEIGHVWVEPYI-VADDVPTG 137

RESULT 11

ABU16806
 ID ABU16806 standard; protein; 352 AA.

XX AC ABU16806;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #2333.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Acinetobacter baumannii.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA20676.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 44730; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 352 AA;

Query Match 9.9%; Score 87; DB 6; Length 352;

Best Local Similarity 25.6%; Pred. No. 2;
 Matches 34; Conservative 20; Mismatches 43; Indels 36; Gaps 6;

QY 37 MVAGRW-----VRDOENVIVKLTGV--YKVSMTPTGTDSVLFNFMPEKRMHGIFFPK 89
 Db 1 MKAARYDRGDIKIEDIPEPEVTPGTGIVKAVCGICGIDL-----HEFMGPIFIPP 53

QY 90 WHEHEPITVCYQNDHIDLMKESREKYETYPKYVVPPEFAITFLKNEGVDSNEV---ISK 146
 Db 54 CGHPHP-----ISGESAPITMGHEFGSVVYAVGVGVDDIIGQHVVV 95

QY 147 APVEGTTDDIRAG.159

Db 96 EPYI-VADDVPTG.107

RESULT 12
 AAB94833
 ID AAB94833 standard; protein; 497 AA.

XX AAB94833;

XX 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:15998.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 15998; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX Sequence 497 AA;

Query Match 9.5%; Score 83.5; DB 4; Length 497;

Best Local Similarity 19.2%; Pred. No. 7.7;
 Matches 32; Conservative 31; Mismatches 49; Indels 55; Gaps 9;

QY 13 YENGWEYETIKNDHTIDYRIHSGWVAGRWVRD--QEVNIVKLTGEG-VYKVSMTPTGTGD 69
 Db 149 FQNGDKYD-----GDWVRDRQGHGVLRCADGSTYKQW----HSD 185

QY 70 VSLNFMPEKRMHGIFFPKWVHEHP-----EITVCYQNDHI 106
 Db 186 V-FSLGSLMAHCSGVTYIGLWINGHPAEQATRIVLGPEVMEVAQSPFSNVNQLLODGH 244

QY 107 DLMKESREK---YETYPKYV-VPEFAITFLKNEGVDSNEVISKAPY 149

Db 245 EIAKSESGRVLQISAGRVYQLSAYSEVNFVKYDR-DNQETLIQTFF 290

RESULT 13

ABP69347

ID ABP69347 standard; protein; 497 AA.

XX ABP69347;

XX 20-JAN-2003 (first entry)

DE Human polypeptide SEQ ID NO 1394.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; neotropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US0005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR N-PSDB; ABZ11564.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 9; SEQ ID NO 1394; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 497 AA;

Query Match 9.5%; Score 83.5; DB 5; Length 497;
Best Local Similarity 19.2%; Pred. No. 7.7;
Matches 32; Conservative 31; Mismatches 49; Indels 55; Gaps 9;

QY 13 YENGWEVEIYKNDHTIDYRIHSGMWAGRWVRD--QEVNIVKLTEG-VYKVSWTEPTGTD 69
DB 149 FQNGDKYD-----GDWDRRRQGHVLRCDAGSTYKGQW----HSD 185

QY 70 VSLNFMPEKRMHGIIPFPKWVHEHP-----BITVCYQNDHI 106
DB 186 V-FSLGSLMAHCSGVTVYGLWINGHPAEQATRVILGPEVMEVAQGSFVSNNVQLLDQH 244

QY 107 DLMKESREK---YETYPKYV-VPFAITFLKNGVNEEIVSKAPY 149
DB 245 EIAKSESGRVLIQISAGVRYVQLSAYSEVNFVKVDR-DNQETLIQTFF 290

RESULT 14
ABP69348
ID ABP69348 standard; protein; 508 AA.
XX
AC ABP69348;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polypeptide SEQ ID NO 1395.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; neotropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US0005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR N-PSDB; ABZ11565.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative, or platelet
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 9; SEQ ID NO 1395; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 508 AA;

Query Match 9.5%; Score 83.5; DB 5; Length 508;
Best Local Similarity 19.2%; Pred. No. 8;
Matches 32; Conservative 31; Mismatches 49; Indels 55; Gaps 9;

QY 13 YENGWEVEIYKNDHTIDYRIHSGMWAGRWVRD--QEVNIVKLTEG-VYKVSWTEPTGTD 69
DB 160 FQNGDKYD-----GDWDRRRQGHVLRCDAGSTYKGQW----HSD 196

QY 70 VSLNFMPEKRMHGIIPFPKWVHEHP-----BITVCYQNDHI 106

Db 197 V-FSGLGMAHCSGVTYGLWINGHPAEQTRIVILGPEVMAQSPFVNVQLLODHG 255
QY 107 DLMKESREK---VETPKYV-VPEFAITFLKNEGVNDNEEVIKAPY 149
Db 256 EIAKESGRVLOISAGRVYVQLSAYSEVNFVKVDR-DNQETLIQTPF 301

RESULT 15

ADT58254 standard; protein; 652 AA.

XX AC ADT58254;

XX DT 13-JAN-2005 (first entry)

XX DE Plant polypeptide, SEQ ID 8331.

XX KW Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.

XX OS Viridiplantae.

XX PN US2004216190-A1.

XX PD 28-OCT-2004.

XX PF 18-DEC-2003; 2003US-00739930.

XX PR 28-APR-2003; 2003US-00424599.

XX PR 28-APR-2003; 2003US-00425115.

XX PA (KOVA/) KOVALIC D K.

XX PI Kovalic DK;

XX DR WPI; 2004-757369/74.

XX PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.

XX PS Claim 2; SEQ ID NO 8331; 14pp; English.

XX CC The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11086). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and

CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.

XX SQ Sequence 652 AA;

Query Match 9.4%; Score 83; DB 8; Length 652;

Best Local Similarity 22.1%; Pred. No. 13;

Matches 46; Conservative 23; Mismatches 63; Indels 76; Gaps 11;

QY 3 NFIGSHMIVTYENG-----WEYEIYIKNDHTI-----DYRTHSG 36

Db 347 NEYGSYDKLKANGTSCILLSCELLDAITKIOWKAPCFSGNGEWIVGASANKGEHRLQIW 406

QY 37 MVAGRWRDQEVNIVKLTEG----VYKVS-TEPTGTDVSLNFMNPNKRMHGIIFFPKW 90

Db 407 NQAGR-----LVQMLEGPKREALIDLAWHHVEPTIATVSVT-----GFVYI--W 447

QY 91 VHEHPETVCYQNDHIDLMKESREKY-----ETYPKYVVPFAEI---TFLK 134

Db 448 AKEHVENWSAFAPDFVEL--EENEYAREDEFDLNPREEQAEKVVIDEADVDIETSEK 505

QY 135 N-----EGVDNEEVIKAPYEGMTDD 155

Db 506 NTVFSDVEDSVDEISYLPALPYDPDSPDE 533

Search completed: October 12, 2006, 05:35:52

Job time : 202 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2006, 05:32:37 ; Search time 52 seconds
(without alignments)
271.008 Million cell updates/sec

Title: US-10-824-581-4

Perfect score: 882

Sequence: 1 MENFGSHMIYTYENGWEYE.....EVISKAPYEGMTDIRAGRL 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5_COMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6_COMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7_COMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H_COMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PCTUS_COMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pep:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	777	88.1	161	2	US-09-235-153-2
2	675	76.5	183	2	US-09-107-532A-6234
3	499	56.6	168	2	US-09-489-039A-10847
4	88.5	10.0	382	2	US-09-328-353-7738
5	80.5	9.1	281	2	US-09-270-767-38987
6	80.5	9.1	281	2	US-09-270-767-54204
7	78.5	8.9	413	2	US-09-248-796A-14672
8	78.5	8.9	1676	2	US-08-487-283A-2
9	78.5	8.9	1680	2	US-09-949-016-9777
10	78	8.8	3135	1	US-08-323-170B-2
11	78	8.8	3135	2	US-08-954-441-2
12	77.5	8.8	391	2	US-09-107-532A-4513
13	77	8.7	937	2	US-09-583-110-3166
14	77	8.7	975	2	US-09-107-433-3032
15	77	8.7	2777	2	US-09-543-681A-6124
16	76.5	8.7	443	2	US-09-248-796A-19547
17	76	8.6	303	2	US-10-360-101-263
18	75.5	8.6	279	2	US-09-704-725-8
19	75.5	8.6	903	2	US-09-270-767-46470
20	75	8.5	397	2	US-09-583-110-4930
21	75	8.5	397	2	US-09-107-433-2897
22	74	8.4	353	2	US-09-984-292-2
23	74	8.4	353	2	US-09-984-292-16
24	74	8.4	359	2	US-09-984-292-1
25	74	8.4	359	2	US-09-984-292-15
26	74	8.4	359	2	US-09-984-292-37

27	74	8.4	835	2	US-09-134-000C-4742	Sequence 4742, Ap
28	73.5	8.3	302	2	US-09-134-000C-4689	Sequence 4689, Ap
29	73.5	8.3	396	2	US-09-107-532A-4739	Sequence 4739, Ap
30	73	8.3	1037	2	US-09-949-016-6723	Sequence 6723, Ap
31	73	8.3	1061	2	US-09-949-016-10165	Sequence 10165, A
32	73	8.3	1297	2	US-09-107-532A-4552	Sequence 4552, Ap
33	72.5	8.2	364	2	US-08-896-537A-3	Sequence 3, Appli
34	72.5	8.2	460	2	US-09-248-796A-17144	Sequence 17144, A
35	72.5	8.2	560	2	US-08-983-045-4	Sequence 4, Appli
36	72	8.2	510	1	US-08-097-829-2	Sequence 2, Appli
37	72	8.2	510	1	US-08-577-403-2	Sequence 2, Appli
38	72	8.2	1348	2	US-09-949-002-517	Sequence 517, App
39	72	8.2	2196	2	US-10-360-101-259	Sequence 259, App
40	72	8.2	2224	2	US-09-054-272-38	Sequence 38, Appl
41	72	8.2	2224	2	US-09-949-002-292	Sequence 292, App
42	71.5	8.1	252	2	US-09-949-016-9408	Sequence 9408, Ap
43	71	8.0	724	2	US-09-562-737-27	Sequence 27, Appl
44	70.5	8.0	303	2	US-09-533-029-6	Sequence 6, Appli
45	70.5	8.0	338	2	US-09-198-452A-1050	Sequence 1050, Ap

ALIGNMENTS

RESULT 1

US-09-235-153-2

; Sequence 2, Application US/09235153

; Patent No. 6703539

; GENERAL INFORMATION:

; APPLICANT: KELLER, WILFRED A.

; APPLICANT: DATLA, RAJU S.S.

; APPLICANT: DONG, JIN-ZHUO

; APPLICANT: GEORGES, FAWZY

; APPLICANT: HUSSAIN, ATTA A. K.

; APPLICANT: SELVARAJ, GOPALAN

; TITLE OF INVENTION: METHODS AND COMPOSITIONS MODIFYING LEVELS OF SECONDARY METABOLIC

; FILE REFERENCE: 73776-127

; CURRENT APPLICATION NUMBER: US/09/235,153

; CURRENT FILING DATE: 1999-01-22

; PRIOR APPLICATION NUMBER: US 60/072156

; PRIOR FILING DATE: 1998-01-22

; PRIOR APPLICATION NUMBER: US 09/012453

; PRIOR FILING DATE: 1998-01-23

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 161

; TYPE: PRT

; ORGANISM: Bacillus pumilus

US-09-235-153-2

Query Match 88.1%; Score 777; DB 2; Length 161;

Best Local Similarity 85.0%; Pred. No. 3.7e-81;

Matches 136; Conservative 16; Mismatches 8; Indels 0; Gaps 0;

QY 1 MENFGSHMIYTYENGWEYEIYKNDHIDYRHSWAGVRDQEVNIVKLTGVYKV 60

Db 1 MDQFVGLHMIYTYENGWEYEIYKNDHIDYRHSWAGVRDQEVNIVKLTGVYKV 60

QY 61 SWTEPTGTVDSLVNFMNPKRMHGIIFPKVWHEHPETVCYQNDHIDLMKESREKTYTP 120

Db 61 SWTEPTGTVDSLVNFMNPKRMHGIIFPKVWHEHPETVCYQNDHIDLMKESREKTYTP 120

QY 121 KYVPEFAEITFLKNQGVNNEEIVISKAPYEGMTDIRAGR 160

Db 121 KYVPEFAEITFLKNQGVNNEEIVISKAPYEGMTDIRAGR 160

RESULT 2

US-09-107-532A-6234

; Sequence 6234, Application US/09107532A

; Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6234:
SEQUENCE CHARACTERISTICS:
LENGTH: 183 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...183
SEQUENCE DESCRIPTION: SEQ ID NO: 6234:
US-09-107-532A-6234

Query Match 76.5%; Score 675; DB 2; Length 183;
Best Local Similarity 70.6%; Pred. No. 2.3e-69;
Matches 113; Conservative 30; Mismatches 17; Indels 0; Gaps 0;
QY 1 MENFIGSHMIYTYENGWEYEIYKNDHTIDYRIHSGMWAGRWVDRDOENVIVKLTEGVYKV 60
Db 14 LEDFIGHFITYDNGWEYEWYAKNDHTVDYRIHSGMWAGRWVDRDOENVIVKLTEGVYKV 73
QY 61 SWTEPTGTDVSLNFMNPKRMHGIIFPKVWVHEHPITVCYQNDHIDLMKESREKYETYP 120
Db 74 TWTEPTGTDVSLNFMNPKRMHGIIFPKVWVHEHPITVCYQNDHIDLMKESREKYETYP 133
QY 121 KYVVEFAEITFLKNEGVNDNEEVISKAPYEGMTDIRAG 160
Db 134 KLVPEFAITFLKNEGVNDNEEVISKAPYEGMTDIRAG 173

RESULT 3
US-09-489-039A-10847
Sequence 10847, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10847
LENGTH: 168
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10847

Query Match 56.6%; Score 499; DB 2; Length 168;
Best Local Similarity 57.9%; Pred. No. 3.3e-49;
Matches 84; Conservative 25; Mismatches 36; Indels 0; Gaps 0;
QY 1 MENFIGSHMIYTYENGWEYEIYKNDHTIDYRIHSGMWAGRWVDRDOENVIVKLTEGVYKV 60
Db 10 LSGFIGKHLVYTYDNGWYEIYVYKNGHTLDYRIHSGIVGNRWVKDQEAIVRVGESIYKI 69
QY 61 SWTEPTGTDVSLNFMNPKRMHGIIFPKVWVHEHPITVCYQNDHIDLMKESREKYETYP 120
Db 70 SWTEPTGTDVSLNFMNPKRMHGIIFPKVWVHEHPITVCYQNDHIDLMKESREKYETYP 129
QY 121 KYVVEFAEITFLKNEGVNDNEEVIS 145
Db 130 TEVIDEFAITFLKNEGVNDNEEVIS 154

RESULT 4
US-09-328-352-7738
Sequence 7738, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7738
LENGTH: 382
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7738

Query Match 10.0%; Score 88.5; DB 2; Length 382;
Best Local Similarity 24.8%; Pred. No. 0.15;
Matches 39; Conservative 23; Mismatches 56; Indels 39; Gaps 7;
QY 13 YENGWEYEIYKNDHTIDYRIHSGMWAGRWVDRDOENVIVKLTEGV--YKVSWTEP 65
Db 10 YFNGRNSLSLKHQHPDLREI---MKAARYDRGDRIEDIPPEVTPGTGVIKVAWCGI 66
QY 66 TGTDVSLNFMNPKRMHGIIFPKVWVHEHPITVCYQNDHIDLMKESREKYETYPKYVVP 125
Db 67 CGTDL-----HEFMEGPIFPPCGHPH-----ISGESAPITMGH 101
QY 126 EFAEITFLKNEGVNDNEEV---ISKAPYEGMTDIRAG 159
Db 102 EFGVYVAVGEGVDDIEIGHVVEPYI-VADDVPTG 137

RESULT 5
US-09-270-767-38987
Sequence 38987, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17

```
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38987
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-38987

Query Match          9.1%; Score 80.5; DB 2; Length 281;
Best Local Similarity 21.0%; Pred. No. 0.78;
Matches 35; Conservative 24; Mismatches 73; Indels 35; Gaps 5;

QY 10 IYTYENGWE-----YEIYIKNDHTIDYRIHSG-----MVAGRWVRDQEVNIVKLTGV 57
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 20 IFTYSNMCFKRYITLYFIMKSKICVYRXRGKQLKVDKIDGLKISXQXALITLKEII 79
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 58 YKVSWTEPTCTDVSLEMPNEKRMHGLIFPKVWHEPEITV--CYQNDHIDLMKES--- 112
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 80 YKLNIYYQTETXSFDLKKSSILNPYIFSLVQIEQSNQILISSCIYSSRSDVISETIXA 139
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 113 -----REKYETYPKVWVP-----EFAEITFLKNEGVNDE 141
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 140 RKLVLTLRVYXTRREYVSGTKYFRPDYDVYKNFKKFFLLXHRITNQ 186
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 6
US-09-270-767-54204
; Sequence 54204, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54204
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54204

Query Match          9.1%; Score 80.5; DB 2; Length 281;
Best Local Similarity 21.0%; Pred. No. 0.78;
Matches 35; Conservative 24; Mismatches 73; Indels 35; Gaps 5;

QY 10 IYTYENGWE-----YEIYIKNDHTIDYRIHSG-----MVAGRWVRDQEVNIVKLTGV 57
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 20 IFTYSNMCFKRYITLYFIMKSKICVYRXRGKQLKVDKIDGLKISXQXALITLKEII 79
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 58 YKVSWTEPTCTDVSLEMPNEKRMHGLIFPKVWHEPEITV--CYQNDHIDLMKES--- 112
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 80 YKLNIYYQTETXSFDLKKSSILNPYIFSLVQIEQSNQILISSCIYSSRSDVISETIXA 139
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 113 -----REKYETYPKVWVP-----EFAEITFLKNEGVNDE 141
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 140 RKLVLTLRVYXTRREYVSGTKYFRPDYDVYKNFKKFFLLXHRITNQ 186
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 7
US-09-248-796A-14672
; Sequence 14672, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
```

```
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14672
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14672

Query Match          8.9%; Score 78.5; DB 2; Length 413;
Best Local Similarity 23.2%; Pred. No. 2.3;
Matches 26; Conservative 17; Mismatches 40; Indels 29; Gaps 4;

QY 20 EIIYKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTGVYKVSWTEPTGTGTVSLNFMPEK 79
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 187 EYVFPNGH-----GDIGGGW-----KASLDNQVLSSVSLRWMLAHA 222
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 80 RMHGLIFPP-----KWVHEHPEITVCYQNDHIDLMKESREKYETYPKVWVPEF 127
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 223 IKYGVRFQPOSISQWSSDHPPISSFLSPNH-DILSFSSKSPDSRSDIVLPEF 273
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 8
US-08-487-283A-2
; Sequence 2, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott
; APPLICANT: Nye, Steven H.
; APPLICANT: Rollins, Scott
; APPLICANT: Rother, Russell P.
; APPLICANT: Springhorn, Jeremy P.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: Thomas, Thomas C.
; APPLICANT: Wilkins, James A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seth A. Fidel
; STREET: 25 Science Park (Alexion)
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06511
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.4Mb storage
; COMPUTER: Macintosh Cetrus 610
; OPERATING SYSTEM: System 7
; SOFTWARE: WordPerfect 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,283A
; FILING DATE: June 7, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,208
; FILING DATE: 02-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Seth A. Fidel.
; REGISTRATION NUMBER: 38,449
; REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)776-1790
; TELEFAX: (203)772-3655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1676 Amino Acids
; TYPE: Amino Acid
```

```

; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE:
; DESCRIPTION: Pro-C5 Polypeptide
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; AUTHORS: Haviland, D.L.
; AUTHORS: Haviland, J.C.
; AUTHORS: Fleischer, D.T.
; AUTHORS: Hunt, A.
; AUTHORS: Wetsel, R.A.
; TITLE: Complete cDNA Sequence of Human
; Patent No. 6355245
; TITLE: Complement Pro-C5
; JOURNAL: Journal of Immunology
; VOLUME: 146
; PAGES: 362-368
; DATE: 1991
;
; US-08-487-283A-2
;
; Query Match 8.9%; Score 78.5; DB 2; Length 1676;
; Best Local Similarity 23.1%; Pred. No. 17;
; Matches 37; Conservative 26; Mismatches 56; Indels 41; Gaps 8;
;
; Qy 9 MIYTYNGWEYIYKNDHTIDYRIHSGMVAGRWVRDQEVNI-----VKLTGEYVKV 60
; Db 117 MPITYDNGF---LFHTDKPV-----YTPDQSVKRVYSLNDDLKPAKRETVL 161
;
; Qy 61 SWTEPTGTDVSLNFMENPKRMHGIIFFPKW-VHEHPE-----ITVCYONDHIDLMKESRE 114
; Db 162 TFDPEGSEVD---MVEIDHIGIISFPDFKIPSNPRYGMWTIKAKYKED-----FSTTGT 214
;
; Qy 115 KYETPKYVVPFEFA-----EITFLKNEGVDNEEVISKAPY 149
; Db 215 AYFEVKEYVLPFHFSVSIEPEYNEFYGNFKNFKNFEITIKARY 254
;
; RESULT 9
; US-09-949-016-9777
; Sequence 9777, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14,755
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9777
; LENGTH: 1680
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-9777
;
; Query Match 8.9%; Score 78.5; DB 2; Length 1680;
; Best Local Similarity 23.1%; Pred. No. 17;
; Matches 37; Conservative 26; Mismatches 56; Indels 41; Gaps 8;
;
; Qy 9 MIYTYNGWEYIYKNDHTIDYRIHSGMVAGRWVRDQEVNI-----VKLTGEYVKV 60
; Db 121 MPITYDNGF---LFHTDKPV-----YTPDQSVKRVYSLNDDLKPAKRETVL 165
;
; Qy 61 SWTEPTGTDVSLNFMENPKRMHGIIFFPKW-VHEHPE-----ITVCYONDHIDLMKESRE 114
; Db 166 TFDPEGSEVD---MVEIDHIGIISFPDFKIPSNPRYGMWTIKAKYKED-----FSTTGT 218
;
; Qy 115 KYETPKYVVPFEFA-----EITFLKNEGVDNEEVISKAPY 149
; Db 219 AYFEVKEYVLPFHFSVSIEPEYNEFYGNFKNFKNFEITIKARY 258
;
; RESULT 10
; US-08-323-170B-2
; Sequence 2, Application US/08323170B
; Patent No. 5733772
; GENERAL INFORMATION:
; APPLICANT: Williamson, Kim C.
; APPLICANT: Kaslow, David C.
; TITLE OF INVENTION: Cloning and Expression of Plasmodium
; TITLE OF INVENTION: Faliciparum Transmission-Blocking Target Antigen, Pfs230
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,170B
; FILING DATE: 13-OCT-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,409
; FILING DATE: 29-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Quine, Jonathan A.
; REGISTRATION NUMBER: P-41,261
; REFERENCE/DOCKET NUMBER: 015280-113100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3135 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-323-170B-2
;
; Query Match 8.8%; Score 78; DB 1; Length 3135;
; Best Local Similarity 22.4%; Pred. No. 49;
; Matches 33; Conservative 28; Mismatches 62; Indels 24; Gaps 6;
;
; Qy 2 ENFTGSHVIYTYNGWEYIYKNDHTIDYRIHSGMVAGRWVRDQEVNI-VKLTEGVYKVS 61
; Db 1292 ENFINPRVKNITFDENVETCNKIENFFNY-----IQIFCPAKDLGIYKNIQWYDI- 1343
;
; Qy 62 WTEPTGTDVSLNFMENPKRMHGIIFFPKWVHEHPEITVCYONDHIDLMKESREKYETVPK 121
; Db 1344 -VKPTRVYPQKXF--NNEELHKLIPNSEMLHKTKEMLILYNEEKVDLL-----HFVVFLLPI 1396
;
; RESULT 11
; US-08-954-441-2
; Sequence 2, Application US/08954441
```



```

; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3166
; LENGTH: 937
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3166

Query Match      8.7%; Score 77; DB 2; Length 937;
Best Local Similarity 23.0%; Pred. No. 11;
Matches 35; Conservative 31; Mismatches 64; Indels 22; Gaps 7;

Qy 7 SHMIYTVENG--WEYEIYIKNDHTID-YRIHSGMVAGRWVRDQEVNIIVKLTEGVYKVSMT 63
Db 566 SRVIYDDHGTWHAAGEAVNDNRQVQDKIHSSTMNKRAQNTSTVVQLNNGDVKLFMR 625
Qy 64 EPTGTVSLNFMNPKRMHGIIFPPKVVHEHPEITVCY-QNDHIDLKESRE----- 114
Db 626 GLTG-DLQV-----ATSKDGGVTWEKDIKRYPOVKDVYVQMSAIHTMHEGKEYIILSNAG 679
Qy 115 --KYETPKYV--VPEFAEITFLKNEGVDNEE 142
Db 680 GPKRENGMVHLARVEENGELTWLKHNPQKGE 711

RESULT 14
US-09-107-433-3032
; Sequence 3032, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTICS
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arintello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3032:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 975 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:

```

```

; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (B) LOCATION 1...975
; SEQUENCE DESCRIPTION: SEQ ID NO: 3032:
US-09-107-433-3032

Query Match      8.7%; Score 77; DB 2; Length 975;
Best Local Similarity 23.0%; Pred. No. 12;
Matches 35; Conservative 31; Mismatches 64; Indels 22; Gaps 7;

Qy 7 SHMIYTVENG--WEYEIYIKNDHTID-YRIHSGMVAGRWVRDQEVNIIVKLTEGVYKVSMT 63
Db 604 SRVIYDDHGTWHAAGEAVNDNRQVQDKIHSSTMNKRAQNTSTVVQLNNGDVKLFMR 663
Qy 64 EPTGTVSLNFMNPKRMHGIIFPPKVVHEHPEITVCY-QNDHIDLKESRE----- 114
Db 664 GLTG-DLQV-----ATSKDGGVTWEKDIKRYPOVKDVYVQMSAIHTMHEGKEYIILSNAG 717
Qy 115 --KYETPKYV--VPEFAEITFLKNEGVDNEE 142
Db 718 GPKRENGMVHLARVEENGELTWLKHNPQKGE 749

RESULT 15
US-09-543-681A-6124
; Sequence 6124, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6124
; LENGTH: 2777
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6124

Query Match      8.7%; Score 77; DB 2; Length 2777;
Best Local Similarity 23.0%; Pred. No. 54;
Matches 34; Conservative 27; Mismatches 51; Indels 36; Gaps 8;

Qy 6 GSHMIYTVENGWEY---EYIKN-DHTIDYRIHSGMVAGRWVRDQEVNIIVKLTEGVYK 60
Db 282 GEHLLYKLKNDWPFPTKMDKSVKYSQTSDDLIFLEKIMGR-----INLVASAENNYNF 335
Qy 61 SWTEPTG---TDVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLKESREK-- 115
Db 336 SHFDPLSGVCYGISLKYL-LEVRNNGIEGKNKYLH-----WLKENINLYKDKKXIS 386
Qy 116 -----YETPKY-----VPEFAEITF 132
Db 387 DKLEKIIFDSVQVEYEVKLKEIKEIF 414

Search completed: October 12, 2006, 05:36:49
Job time : 53 secs

```


GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM protein - protein search, using sw model
 Run on: October 12, 2006, 05:35:51 | Search time 85 Seconds
 (without alignments)
 877.383 Million cell updates/sec
 Title: US-10-824-581-4
 Perfect score: 882
 Sequence: 1 MENFIGSHMIYTYENGWEYE.....EVISKAPYEGMTDDIRAGRL 161
 Scoring table: BLOSUM62
 Gapop 10.0, Gapext 0.5
 Searched: 2097797 seqs, 463214958 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database: Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	882	100.0	161	4	US-10-439-478-6
2	882	100.0	161	5	US-10-824-581-4
3	777	88.1	161	4	US-10-253-971-2
4	666.5	75.6	174	4	US-10-439-478-4
5	666.5	75.6	174	5	US-10-824-581-2
6	142	16.1	176	4	US-10-425-115-266715
7	91.5	10.4	1574	4	US-10-369-493-147
8	87	9.9	352	4	US-10-282-122A-44730
9	83.5	9.5	827	5	US-10-732-923-18311
10	83	9.4	366	4	US-10-425-115-265879
11	83	9.4	652	5	US-10-739-930-8331
12	83	9.4	827	5	US-10-732-923-18337
13	81	9.2	233	5	US-10-732-923-18890
14	81	9.2	233	6	US-11-045-004-1741
15	80.5	9.1	407	6	US-11-079-463-8975
16	80.5	9.1	410	5	US-10-467-657-2312
17	80	9.1	800	4	US-10-474-776-363
18	80	9.1	800	5	US-10-472-928-3488
19	80	9.1	1168	6	US-11-079-463-9340
20	80	9.1	1368	4	US-10-369-493-11554
21	79.5	9.0	315	4	US-10-437-963-135413
22	79	9.0	251	5	US-10-739-930-6394
23	79	9.0	354	4	US-10-282-122A-64997
24	79	9.0	740	5	US-10-450-763-42199
25	78.5	8.9	1251	3	US-09-778-927A-58
26	78.5	8.9	1602	3	US-09-778-927A-59
27	78.5	8.9	1676	5	US-10-719-993-785

28 78.5 8.9 1676 6 US-11-058-134A-102 Sequence 102, App
 29 78.5 8.9 1680 4 US-10-724-662-2 Sequence 2, Appli
 30 78.5 8.9 1680 5 US-10-989-891-142 Sequence 142, App
 31 78 8.8 233 5 US-10-732-923-18889 Sequence 18889, A
 32 78 8.8 452 6 US-11-097-143-2907 Sequence 2907, Ap
 33 77 8.7 354 6 US-11-087-099-12051 Sequence 12051, A
 34 77 8.7 461 4 US-10-425-114-56187 Sequence 56187, A
 35 77 8.7 975 5 US-10-617-320-3032 Sequence 3032, Ap
 36 76.5 8.7 315 5 US-10-764-260-27 Sequence 27, Appl
 37 76.5 8.7 398 4 US-10-369-493-10194 Sequence 10194, A
 38 76.5 8.7 752 5 US-10-732-923-12812 Sequence 12812, A
 39 76.5 8.7 752 6 US-11-087-099-11284 Sequence 11284, A
 40 76 8.6 303 4 US-10-360-101-263 Sequence 263, App
 41 76 8.6 354 4 US-10-282-122A-65796 Sequence 65796, A
 42 76 8.6 354 6 US-11-087-099-6537 Sequence 6537, Ap
 43 76 8.6 400 6 US-11-087-099-5995 Sequence 5995, Ap
 44 76 8.6 657 5 US-10-732-923-13661 Sequence 13661, A
 45 75.5 8.6 513 5 US-10-510-408-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
 US-10-439-478-6
 ; Sequence 6, Application US/10439478
 ; Publication No. US20040018600A1
 ; GENERAL INFORMATION:
 ; APPLICANT: E.I. du Pont de Nemours and Co.
 ; APPLICANT: Ben-Bassat, Arle
 ; APPLICANT: Oi, Wei Wei
 ; APPLICANT: Sariaslani, Sima
 ; APPLICANT: Tang, Xiao-Song
 ; APPLICANT: Vanelli, Todd
 ; TITLE OF INVENTION: Microbial Conversion of Glucose to Para-Hydroxystyrene
 ; FILE REFERENCE: CL1912
 ; CURRENT APPLICATION NUMBER: US/10/439,478
 ; CURRENT FILING DATE: 2003-05-16
 ; PRIOR APPLICATION NUMBER: 60/383450
 ; PRIOR FILING DATE: 2002-05-23
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 6
 ; LENGTH: 161
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilis
 ; US-10-439-478-6
 Query Match 100.0%; Score 882; DB 4; Length 161;
 Best Local Similarity 100.0%; Pred No. 6.6e-80;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MENFIGSHMIYTYENGWEYEIYKNDHTIDYRIHSGMVGWRVDRQEVNIVKLTEGYKV 60
 Db 1 MENFIGSHMIYTYENGWEYEIYKNDHTIDYRIHSGMVGWRVDRQEVNIVKLTEGYKV 60
 QY 61 SWTEPTGTVSLNFMNPNKXRMGIIFFPKVWHEHPITVCYQNDHIDLMKESREKYETYP 120
 Db 61 SWTEPTGTVSLNFMNPNKXRMGIIFFPKVWHEHPITVCYQNDHIDLMKESREKYETYP 120
 QY 121 KYVWPEFASITFLKNGVDNEEVISKAPYEGMTDDIRAGRL 161
 Db 121 KYVWPEFASITFLKNGVDNEEVISKAPYEGMTDDIRAGRL 161
 RESULT 2
 US-10-824-581-4
 ; Sequence 4, Application US/10824581
 ; Publication No. US20040248267A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haynie, Sharon L
 ; APPLICANT: Ben-Bassat, Arle
 ; APPLICANT: Lowe, David J.

; APPLICANT: Huang, Lisa L.
 ; TITLE OF INVENTION: A Method for Producing Para-Hydroxystyrene by Biocatalytic
 ; TITLE OF INVENTION: Decarboxylation of Hydroxycinnamic Acid in a Biphasic Reaction
 ; TITLE OF INVENTION: Medium
 ; FILE REFERENCE: CL 2371
 ; CURRENT APPLICATION NUMBER: US/10/824,581
 ; CURRENT FILING DATE: 2004-04-14
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4
 ; LENGTH: 161
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilis
 ; ORGANISM: Bacillus subtilis
 ; US-10-824-581-4

Query Match 100.0%; Score 882; DB 5; Length 161;
 Best Local Similarity 100.0%; Pred. No. 6.6e-80;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVDRDQEVNIVKLTEGYKV 60
 DB 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVDRDQEVNIVKLTEGYKV 60
 QY 61 SWTEPTGTDVSLNFMPEKRMHGIIFPPKWHHEPEITVCYQNDHIDLMKESREKYETYP 120
 DB 61 SWTEPTGTDVSLNFMPEKRMHGIIFPPKWHHEPEITVCYQNDHIDLMKESREKYETYP 120
 QY 121 KYVPEFAEITFLKNEGVNNEEVISKAPYEGMTDDIRAGRL 161
 DB 121 KYVPEFAEITFLKNEGVNNEEVISKAPYEGMTDDIRAGRL 161

RESULT 3
 US-10-253-971-2
 ; Sequence 2, Application US/10253971
 ; Publication No. US20030070192A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GEORGES, FAWZY
 ; APPLICANT: DONG, JIN-ZHUO
 ; APPLICANT: KELLER, WILF
 ; APPLICANT: HUSSAIN, ATTA A. K.
 ; APPLICANT: SELVARAJ, GOPALAN
 ; APPLICANT: DATLA, RAJU
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING SECONDARY PLANT
 ; FILE REFERENCE: 73776-159D
 ; CURRENT APPLICATION NUMBER: US/10/253,971
 ; CURRENT FILING DATE: 2002-09-25
 ; PRIOR APPLICATION NUMBER: US 60/072156
 ; PRIOR FILING DATE: 1998-01-22
 ; PRIOR APPLICATION NUMBER: US 09/012453
 ; PRIOR FILING DATE: 1998-01-23
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 161
 ; TYPE: PRT
 ; ORGANISM: Bacillus pumilus
 ; US-10-253-971-2

Query Match 88.1%; Score 777; DB 4; Length 161;
 Best Local Similarity 85.0%; Pred. No. 2e-69;
 Matches 136; Conservative 16; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVDRDQEVNIVKLTEGYKV 60
 DB 1 MDQVGLHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVDRDQEVNIVKLTEGYKV 60
 QY 61 SWTEPTGTDVSLNFMPEKRMHGIIFPPKWHHEPEITVCYQNDHIDLMKESREKYETYP 120
 DB 61 SWTEPTGTDVSLNFMPEKRMHGIIFPPKWHHEPEITVCYQNDHIDLMKESREKYETYP 120
 QY 121 KYVPEFAEITFLKNEGVNNEEVISKAPYEGMTDDIRAGRL 160

; APPLICANT: Huang, Lisa L.
 ; TITLE OF INVENTION: A Method for Producing Para-Hydroxystyrene by Biocatalytic
 ; TITLE OF INVENTION: Decarboxylation of Hydroxycinnamic Acid in a Biphasic Reaction
 ; TITLE OF INVENTION: Medium
 ; FILE REFERENCE: CL 2371
 ; CURRENT APPLICATION NUMBER: US/10/824,581
 ; CURRENT FILING DATE: 2004-04-14
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4
 ; LENGTH: 161
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilis
 ; ORGANISM: Bacillus subtilis
 ; US-10-824-581-4

Query Match 100.0%; Score 882; DB 5; Length 161;
 Best Local Similarity 100.0%; Pred. No. 6.6e-80;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVDRDQEVNIVKLTEGYKV 60
 DB 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVDRDQEVNIVKLTEGYKV 60
 QY 61 SWTEPTGTDVSLNFMPEKRMHGIIFPPKWHHEPEITVCYQNDHIDLMKESREKYETYP 120
 DB 61 SWTEPTGTDVSLNFMPEKRMHGIIFPPKWHHEPEITVCYQNDHIDLMKESREKYETYP 120
 QY 121 KYVPEFAEITFLKNEGVNNEEVISKAPYEGMTDDIRAGRL 161
 DB 121 KYVPEFAEITFLKNEGVNNEEVISKAPYEGMTDDIRAGRL 161

RESULT 4
 US-10-439-478-4
 ; Sequence 4, Application US/10439478
 ; Publication No. US20040018600A1
 ; GENERAL INFORMATION:
 ; APPLICANT: E.I. du Pont de Nemours and Co.
 ; APPLICANT: Ben-Bassat, Arie
 ; APPLICANT: Qi, Wei Wei
 ; APPLICANT: Sariaslani, Sima
 ; APPLICANT: Tang, Xiao-Song
 ; APPLICANT: Vanelli, Todd
 ; TITLE OF INVENTION: Microbial Conversion of Glucose to Para-Hydroxystyrene
 ; FILE REFERENCE: CL1912
 ; CURRENT APPLICATION NUMBER: US/10/439,478
 ; CURRENT FILING DATE: 2003-05-16
 ; PRIOR APPLICATION NUMBER: 60/383450
 ; PRIOR FILING DATE: 2002-05-23
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 4
 ; LENGTH: 174
 ; TYPE: PRT
 ; ORGANISM: Lactobacillus plantarum
 ; US-10-439-478-4

Query Match 75.6%; Score 666.5; DB 4; Length 174;
 Best Local Similarity 71.9%; Pred. No. 2.5e-58;
 Matches 115; Conservative 26; Mismatches 18; Indels 1; Gaps 1;
 QY 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVDRDQEVNIVKLTEGYKV 60
 DB 8 LDDFLGTHFIYTYDNGWEYEWYAKNDHTVDYRIHSGMVAGRWVDRDQADIVMLTEGIYKI 67
 QY 61 SWTEPTGTDVSLNFMPEKRMHGIIFPPKWHHEPEITVCYQNDHIDLMKESREKYETYP 120
 DB 68 SWTEPTGTDVSLNFMPEKRMHGIIFPPKWHHEPEITVCYQNDHIDLMKESREKYETYP 127
 QY 121 KYVPEFAEITFLKNEGVNNEEVISKAPYEGMTDDIRAGRL 160
 DB 128 KLVVPEFANITYM-GEQONNEDVISEAPYKEMPNDIRNGK 166

Query Match 75.6%; Score 666.5; DB 5; Length 174;
 Best Local Similarity 71.9%; Pred. No. 2.5e-58;
 Matches 115; Conservative 26; Mismatches 18; Indels 1; Gaps 1;

Db 1 MKAARFYDRGDIRIEDIPEVPTGVIGIKVAMCGICGTDL-----HEFMGPIPIPP 53
Qy 90 WVEHPHPTVCYQNDHIDLMKESREKYEYKPVVPEFAITFLKNEGVNNEV---ISK 146
Db 54 CGHPHP-----ISGESAPITMGHEFGVWYAVGSGVDDIBIGQHVW 95
Qy 147 AYEGMTDDIRAG 159
Db 96 EPYI-VADDVPTG 107

RESULT 9

US-10-732-923-18311

; Sequence 18311, Application US/10732923

; Publication No. US20050108791A1

; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D

; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

; FILE REFERENCE: 38-15(52796)C

; CURRENT APPLICATION NUMBER: US/10/732,923

; CURRENT FILING DATE: 2003-12-10

; PRIOR APPLICATION NUMBER: 10/310,154

; PRIOR FILING DATE: 2002-12-04

; NUMBER OF SEQ ID NOS: 24149

; SEQ ID NO 18311

; LENGTH: 827

; TYPE: PRT

; ORGANISM: Fusobacterium nucleatum subsp. nucleatum ATCC 25586

US-10-732-923-18311

Query Match 9.5%; Score 83.5; DB 5; Length 827;
Best Local Similarity 27.3%; Pred. No. 32; Indels 37; Gaps 10;
Matches 41; Conservative 22; Mismatches 50;

Qy 19 YEI-----YIKNDHTIDYRIHS-----GMVAGRWVRDQEV-NIVKLTGEGYKVSMT 63
Db 231 YEISKLDNSYLVENEEOQLDSIFRGKEVVGITAGASTPEETIMNIEKKVGIYKMSV 290
Qy 64 EPTGTVSL---NMPN-EKRMGIIIFPKVWHHPHPTVCYQN-DHIDLMKES---REK 115
Db 291 NENQNEPSLMEELPNOEKREKGV-----ESMD-----QNFSLDVPGETAVRVR 338
Qy 116 YETPKVVPPEFAR--ITFLKNEGVNNEV 143
Db 339 TDELKDYKGVTVTEVLITGLSEEDDDQEI 368

RESULT 10

US-10-425-115-265879

; Sequence 265879, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 265879

; LENGTH: 366

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_174086C.1.pap

US-10-425-115-265879

Query Match 9.4%; Score 83; DB 4; Length 366;
Best Local Similarity 22.1%; Pred. No. 13;
Matches 46; Conservative 23; Mismatches 63; Indels 76; Gaps 11;

Qy 3 NFIGSHMIYTYENG-----WEYBIYIKNDHTI-----DYRIHSG 36
Db 61 NEYGSYDKKANGTSCLLISCELLDAITKIQWAPCFSGNGEWIVGASANKGEHRLQIW 120
Qy 37 MVAGRWRDQEVNIVKLTG-----VYKVSW--TEPTGTDVSLNFMNPKRMHGIIFPKW 90
Db 121 NQAGR-----LVKMLEGPKREALDLAWHPVEPTIATVSVT-----GFVYI--W 161
Qy 91 VHEHPHPTVCYQNDHIDLMKESREK-----ETYPKYVVPEFAEI---TFLK 134
Db 162 AKEHVENWSAFAPDFVEL--EENEYAEAREDEFDLNPREEQAEKWIDEDADVDTETSEK 219
Qy 135 N-----EGVDNEEVISKAPYEGMTDD 155
Db 220 NTVFSDVEDSVDEISYLPALPYDPSDE 247

RESULT 11

US-10-739-930-8331

; Sequence 8331, Application US/10739930

; Publication No. US20040216190A1

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

; FILE REFERENCE: 38-21(53377)B

; CURRENT APPLICATION NUMBER: US/10/739,930

; CURRENT FILING DATE: 2003-12-18

; NUMBER OF SEQ ID NOS: 11088

; SEQ ID NO 8331

; LENGTH: 652

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C49976_1.p

US-10-739-930-8331

Query Match 9.4%; Score 83; DB 5; Length 652;
Best Local Similarity 22.1%; Pred. No. 27; Indels 76; Gaps 11;
Matches 46; Conservative 23; Mismatches 63;

Qy 3 NFIGSHMIYTYENG-----WEYBIYIKNDHTI-----DYRIHSG 36
Db 347 NEYGSYDKKANGTSCLLISCELLDAITKIQWAPCFSGNGEWIVGASANKGEHRLQIW 406
Qy 37 MVAGRWRDQEVNIVKLTG-----VYKVSW--TEPTGTDVSLNFMNPKRMHGIIFPKW 90
Db 407 NQAGR-----LVKMLEGPKREALDLAWHPVEPTIATVSVT-----GFVYI--W 447
Qy 91 VHEHPHPTVCYQNDHIDLMKESREK-----ETYPKYVVPEFAEI---TFLK 134
Db 448 AKEHVENWSAFAPDFVEL--EENEYAEAREDEFDLNPREEQAEKWIDEDADVDTETSEK 505
Qy 135 N-----EGVDNEEVISKAPYEGMTDD 155
Db 506 NTVFSDVEDSVDEISYLPALPYDPSDE 533

RESULT 12

US-10-732-923-18337

; Sequence 18337, Application US/10732923

; Publication No. US20050108791A1

; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D

; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

; FILE REFERENCE: 38-15(52796)C

; CURRENT APPLICATION NUMBER: US/10/732,923

; CURRENT FILING DATE: 2003-12-10

; PRIOR APPLICATION NUMBER: 10/310,154

; PRIOR FILING DATE: 2002-12-04

; NUMBER OF SEQ ID NOS: 24149

; SEQ ID NO 18337

```

; LENGTH: 827
; TYPE: PRT
; ORGANISM: Fusobacterium nucleatum subsp. vincentii ATCC 49256
US-10-732-923-18337

Query Match          9.4%; Score 83; DB 5; Length 827;
Best Local Similarity 28.3%; Pred. No. 36;
Matches 41; Conservative 22; Mismatches 52; Indels 30; Gaps 10;

Qy 15 NGWEYEIYKNDHTIDYRIHS-----GMVAGRWRDQEV-NIVKLTGVYKVSWTPTGT 68
Db 238 NGNSY--LVNEEQDLTIFRGKEWGITAGASTPEETIMNIEKKIRGIYKMSNANQN 295
Qy 69 DVSL---NFWPN-EKRMHGIFFPKWVHEPITVCYQN-DHIDLKES---REKYETYP 120
Db 296 EFSLMLEEFNPQKRVGVGVI-----ESMD-----QNSYLDVPGERTAVRVTDELK 343
Qy 121 KYVVPFAE--ITFLKNEGVNDEEV 143
Db 344 GYKVGDTVEVLITGLSEEDDQEVI 368

RESULT 13
US-10-732-923-18890
; Sequence 18890, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Egerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 18890
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Listeria monocytogenes EGD-e
US-10-732-923-18890

Query Match          9.2%; Score 81; DB 5; Length 233;
Best Local Similarity 23.7%; Pred. No. 11;
Matches 42; Conservative 24; Mismatches 69; Indels 42; Gaps 7;

Qy 11 YTYENGWEYIYKNDHTIDYRIHSGMVAGRWRDQEVNIVKLTGVYKVSWTE----- 64
Db 10 YDEKTPGEMWGHICSDFEI---AHTGK-AQSPVDIEQADVVKLKPSTMKFYKETDYTIR 65
Qy 65 -----PTGTDVSLNF-----MPNEKRMHGIFFPKW--VHEHPEITVC 100
Db 66 RIQSVHVFPHDKXEQGLRFNGEYPLVSLFHAHIPAEHLDDGYIYPIEWHFVHEKPDGTTL 125
Qy 101 YQNDHIDLKESREKYETYPKYVVPEFA-----EITFLKNEGVNDEEVISKAPYEG 151
Db 126 VMSAWMDIDNTNNIEFKNLPTYPEVFPADPETEREITLDVNEFMPEERVF--YTYQG 180

RESULT 14
US-11-045-004-1741
; Sequence 1741, Application US/11045004
; Publication No. US20060078901A1
; GENERAL INFORMATION:
; APPLICANT: BUCHRIESER, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
; APPLICANT: GLASER, PHILIPPE

```

```

; APPLICANT: KUNT, FRANK
; APPLICANT: COSSART, PASCALE
; APPLICANT: DANIELS, JUSTIN
; APPLICANT: GOEBEL, WERNER
; APPLICANT: KREFT, JURGEN
; APPLICANT: KUHN, MICHAEL
; APPLICANT: NG, EVA
; APPLICANT: VAZQUEZ-BOLAND, ANTONIO
; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
; APPLICANT: GARRIDO-GARCIA, PATRICIA
; APPLICANT: TIERREZ-MARTINEZ, ALBERTO
; APPLICANT: AMEND, ALEXANDRA
; APPLICANT: CHAKRABORTY, TRINAD
; APPLICANT: DOMANN, EUGEN
; APPLICANT: HAIN, THORSTEN
; APPLICANT: BERCHE, PATRICK
; APPLICANT: DURANT, LIONEL
; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
; APPLICANT: BAQUERO, FERNANDO
; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
; APPLICANT: GOMEZ-LOPEZ, NURIA
; APPLICANT: MADUENIO, ENCARNIA
; APPLICANT: WEHLAND, JURGEN
; APPLICANT: KARST, UWE
; APPLICANT: ENTIAN, KARL-DIETER
; APPLICANT: HAUF, JORG
; APPLICANT: VOSS, HAMUT
; APPLICANT: ROSE, MATTHIAS
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
; FILE REFERENCE: 05394.0018-02
; CURRENT APPLICATION NUMBER: US/11/045,004
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 10/637,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: 10/257,023
; PRIOR FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: PCT/FR01/01118
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: FR 00/04,629
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 2854
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1741
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-11-045-004-1741

Query Match          9.2%; Score 81; DB 6; Length 233;
Best Local Similarity 23.7%; Pred. No. 11;
Matches 42; Conservative 24; Mismatches 69; Indels 42; Gaps 7;

Qy 11 YTYENGWEYIYKNDHTIDYRIHSGMVAGRWRDQEVNIVKLTGVYKVSWTE----- 64
Db 10 YDEKTPGEMWGHICSDFEI---AHTGK-AQSPVDIEQADVVKLKPSTMKFYKETDYTIR 65
Qy 65 -----PTGTDVSLNF-----MPNEKRMHGIFFPKW--VHEHPEITVC 100
Db 66 RIQSVHVFPHDKXEQGLRFNGEYPLVSLFHAHIPAEHLDDGYIYPIEWHFVHEKPDGTTL 125
Qy 101 YQNDHIDLKESREKYETYPKYVVPEFA-----EITFLKNEGVNDEEVISKAPYEG 151
Db 126 VMSAWMDIDNTNNIEFKNLPTYPEVFPADPETEREITLDVNEFMPEERVF--YTYQG 180

RESULT 15
US-11-079-463-8975
; Sequence 8975, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton

```

Search completed: October 12, 2006, 05:37:25
Job time : 87 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2006, 05:36:12 ; Search time 16 Seconds
(without alignments)
799.016 Million cell updates/sec

Title: US-10-824-581-4

Perfect score: 882

Sequence: 1 MENFIGSHMITYENGWEYE.....EVISKAPYEGMTDIRAGRL 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 295242 seqs, 79405279 residues

Total number of hits satisfying chosen parameters: 295242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /EMC Celerra_SIDS3/prodata/2/pubpaa/US09_NEW_PUB.psp:
- 2: /EMC Celerra_SIDS3/prodata/2/pubpaa/US06_NEW_PUB.psp:
- 3: /EMC Celerra_SIDS3/prodata/2/pubpaa/US07_NEW_PUB.psp:
- 4: /EMC Celerra_SIDS3/prodata/2/pubpaa/US08_NEW_PUB.psp:
- 5: /EMC Celerra_SIDS3/prodata/2/pubpaa/PCT_NEW_PUB.psp:
- 6: /EMC Celerra_SIDS3/prodata/2/pubpaa/US10_NEW_PUB.psp:
- 7: /EMC Celerra_SIDS3/prodata/2/pubpaa/US11_NEW_PUB.psp:
- 8: /EMC Celerra_SIDS3/prodata/2/pubpaa/US60_NEW_PUB.psp:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	9.4	367	7	US-11-330-403-7428
2	83	9.4	652	7	US-11-330-403-12464
3	78.5	8.9	1676	7	US-11-318-227-102
4	78	8.8	1455	7	US-11-056-3558-81256
5	78	8.8	1510	7	US-11-056-3558-81255
6	78	8.8	1523	7	US-11-056-3558-81254
7	77	8.7	354	7	US-11-330-403-6834
8	76.5	8.7	752	7	US-11-330-403-4821
9	76.5	8.7	1370	7	US-11-434-137-4210
10	76.5	8.7	1370	7	US-11-434-184-4210
11	76.5	8.7	1370	7	US-11-434-199-4210
12	76.5	8.7	1370	7	US-11-434-203-4210
13	75	8.5	1077	6	US-10-449-902-47193
14	74.5	8.4	287	7	US-11-056-3558-49400
15	74.5	8.4	422	7	US-11-056-3558-16491
16	74.5	8.4	454	7	US-11-056-3558-16490
17	74.5	8.4	471	7	US-11-056-3558-49399
18	74.5	8.4	970	7	US-11-301-764-120
19	73.5	8.3	2303	7	US-11-335-891-31
20	73	8.3	456	7	US-11-056-3558-18682
21	73	8.3	458	7	US-11-056-3558-18681
22	73	8.3	467	7	US-11-056-3558-18680
23	72.5	8.2	298	6	US-10-449-902-40541
24	72.5	8.2	364	6	US-10-511-937-2927
25	72.5	8.2	364	7	US-11-246-999-92

26	72	8.2	480	7	US-11-056-3558-17266	Sequence 17266, A
27	72	8.2	510	7	US-11-330-403-10535	Sequence 10535, A
28	72	8.2	539	6	US-10-449-902-53983	Sequence 53983, A
29	71.5	8.1	479	7	US-11-056-3558-34844	Sequence 34844, A
30	71.5	8.1	479	7	US-11-056-3558-80655	Sequence 80655, A
31	71.5	8.1	481	7	US-11-056-3558-81099	Sequence 81099, A
32	71.5	8.1	486	7	US-11-056-3558-34843	Sequence 34843, A
33	71.5	8.1	486	7	US-11-056-3558-80654	Sequence 80654, A
34	71.5	8.1	534	7	US-11-056-3558-80654	Sequence 80654, A
35	71.5	8.1	987	6	US-10-449-902-45475	Sequence 45475, A
36	71	8.0	312	6	US-10-953-349-19389	Sequence 19389, A
37	71	8.0	359	6	US-10-953-349-19388	Sequence 19388, A
38	71	8.0	370	6	US-10-953-349-19387	Sequence 19387, A
39	71	8.0	724	7	US-11-443-904-27	Sequence 27, Appl
40	71	8.0	948	6	US-10-449-902-53220	Sequence 53220, A
41	70.5	8.0	211	7	US-11-434-137-3470	Sequence 3470, Ap
42	70.5	8.0	211	7	US-11-434-184-3470	Sequence 3470, Ap
43	70.5	8.0	211	7	US-11-434-199-3470	Sequence 3470, Ap
44	70.5	8.0	211	7	US-11-434-203-3470	Sequence 3470, Ap
45	70.5	8.0	291	6	US-10-953-349-4472	Sequence 4472, Ap

ALIGNMENTS

RESULT 1
US-11-330-403-7428
; Sequence 7428, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 7428
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(367)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-330-403-7428

Query Match 9.4%; Score 83; DB 7; Length 367;
Best Local Similarity 22.1%; Pred. No. 2.3;
Matches 46; Conservative 23; Mismatches 63; Indels 76; Gaps 11;

QY	3	NFIGSHMITYENG	-----WEYEIYIKNDHTI-----DYRTHSG 36
DB	61	NEYGSHYDKLKGANGTSCLLISCELLDAITKIQWAPCFSGNGEWIVGASANKGHEHRLQIW 120	-----VYKVSW--TEPTGTDVSLNFMFNEKRMHGIFFPKW 90
QY	37	MVAGRWVRDOENVIKLITEG	-----LVKLEGPKEALIDLAWHPVEPTIATVSVT-----GFPYI--W 161
DB	121	NOAGR	-----LVKLEGPKEALIDLAWHPVEPTIATVSVT-----GFPYI--W 161
QY	91	VHEHPITVCYQNDHIDLKMKESREKY	-----ETYPKYVVPFPAEI---TFLK 134
DB	162	AKEHVENWSAFAPDFVEL	-----EENEYAEAREDEFDLNPREQABKVVIDEADVDIETSEK 219
QY	135	N-----EGVDNEEVISKAPYEGMTDD 155	
DB	220	NTVFSVDSDVDEISYLPALPYPDSPDE 247	

RESULT 2
US-11-330-403-12464
; Sequence 12464, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.

;; TITLE OF INVENTION: Genes and Uses for Plant Improvement

;; FILE REFERENCE: 38-21(53629)B

;; CURRENT APPLICATION NUMBER: US/11/330,403

;; CURRENT FILING DATE: 2006-01-12

;; NUMBER OF SEQ ID NOS: 19250

;; SEQ ID NO 12464

;; LENGTH: 652

;; TYPE: PRT

;; ORGANISM: Zea mays

US-11-330-403-12464

Query Match 9.4%; Score 83; DB 7; Length 652;

Best Local Similarity 22.1%; Pred. No. 4.7;

Matches 46; Conservative 23; Mismatches 63; Indels 76; Gaps 11;

Qy 3 NFIGSHMIYTYENG-----WEYEIYKNDHTI-----DYRIHSG 36

Db 347 NEYSGHYDKLANGTSCLLSCELLDAITKIQKAPCFSGNGEWLVGASANKGEHRLQIW 406

Qy 37 MVAGRWRDQEVNIIVKLTEG-----VYKVSW--TEPTGTDVSLNFMPEKRMHGIIFPKW 90

Db 407 NQAGR-----LVKMLEGKEALIDIAWHVETIATVSVT-----GFYIT--W 447

Qy 91 VHEHPEITVQYONHDIDLMKESREKY-----ETYPKYVVPPEAEI---TFLK 134

Db 448 AKEHVENWSAFDPFVEL--SENSEYAREDEFDLNPREEQAEKVVIDEDADVDETSEK 505

Qy 135 N-----EGVDNEEIVSKAPYEGMTDD 155

Db 506 NTVFSDVEDSVDEISYLPAPYDPSDE 533

RESULT 3

US-11-318-227-102

;; Sequence 102, Application US/11318227

;; Publication No. US20060105980A1

;; GENERAL INFORMATION:

;; APPLICANT: Benedict, Claude

;; APPLICANT: Diener, John

;; APPLICANT: Epstein, David

;; APPLICANT: Grate, Dilara

;; APPLICANT: Keene, Sara Chesworth

;; APPLICANT: Kurz, Jeffrey

;; APPLICANT: Kurz, Markus

;; APPLICANT: McCauley, Thomas Green

;; APPLICANT: Rottman, James

;; APPLICANT: Thompson, Kristin

;; APPLICANT: Wilson, Charles

;; APPLICANT: Zoltoski, Anna J.

;; TITLE OF INVENTION: Aptamer Therapeutics Useful in the Treatment of Complement

;; FILE REFERENCE: 23239-576 CON

;; CURRENT APPLICATION NUMBER: US/11/318,227

;; CURRENT FILING DATE: 2005-12-22

;; PRIOR APPLICATION NUMBER: 11/058,134

;; PRIOR FILING DATE: 2005-02-14

;; PRIOR APPLICATION NUMBER: 60/608,048

;; PRIOR FILING DATE: 2004-09-07

;; PRIOR APPLICATION NUMBER: 60/581,685

;; PRIOR FILING DATE: 2004-06-21

;; PRIOR APPLICATION NUMBER: 60/547,747

;; PRIOR FILING DATE: 2004-02-25

;; PRIOR APPLICATION NUMBER: 60/544,542

;; PRIOR FILING DATE: 2004-02-12

;; NUMBER OF SEQ ID NOS: 102

;; SOFTWARE: PatentIn version 3.3

;; SEQ ID NO 102

;; LENGTH: 1676

;; TYPE: PRT

;; ORGANISM: artificial

;; FEATURE:

;; OTHER INFORMATION: synthetic C5

US-11-318-227-102

Query Match 8.9%; Score 78.5; DB 7; Length 1676;

Best Local Similarity 23.1%; Pred. No. 44;

Matches 37; Conservative 26; Mismatches 56; Indels 41; Gaps 8;

Qy 9 MIYTVENGWEYEIYKNDHTIDYRIHSGMVAGRWVRDQEVNI-----VKLTGEVYKV 60

Db 117 MPITYDNGF---LFIHTDKPV-----YTPDQSVKRVYSLNDDLKPAKRETVL 161

Qy 61 SWTEPTGTDVSLNFMPEKRMHGIIFPKW-VHEHPE-----ITVCYQNDHIDLMKESRE 114

Db 162 TFDIDPEGSEVD---MVEEIDHIGIISFPDPFKIPSNPRYGMWTIKAKYKED----FSTTGT 214

Qy 115 KYTYPKYVVPPEA-----EITFLKNEGVNDNEEVISKAPY 149

Db 215 AYFEVKEYVLPHPFSVSIPEYINFIGYKNFKNFETITKARY 254

RESULT 4

US-11-056-355B-81256

;; Sequence 81256, Application US/11056355B

;; Publication No. US20060150283A1

;; GENERAL INFORMATION:

;; APPLICANT: Brover, Vyacheslav

;; APPLICANT: Alexandrov, Nikolai

;; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding

;; TITLE OF INVENTION: Polypeptides Encoded Thereby

;; FILE REFERENCE: 2750-1590PUS2

;; CURRENT APPLICATION NUMBER: US/11/056,355B

;; CURRENT FILING DATE: 2005-02-14

;; PRIOR APPLICATION NUMBER: 60/544,190

;; PRIOR FILING DATE: 2004-02-13

;; NUMBER OF SEQ ID NOS: 119966

;; SEQ ID NO 81256

;; LENGTH: 1455

;; TYPE: prt

;; ORGANISM: Arabidopsis thaliana

;; FEATURE:

;; NAME/KEY: peptide

;; LOCATION: (1)..(1455)

;; OTHER INFORMATION: Ceres Seq. ID no. 12659032

US-11-056-355B-81256

Query Match 8.8%; Score 78; DB 7; Length 1455;

Best Local Similarity 21.7%; Pred. No. 41;

Matches 33; Conservative 20; Mismatches 43; Indels 56; Gaps 6;

Qy 21 IYIKNDHTIDYRIHSGMVAG--RWVRDQEVNIIVKLTEGVYKVSMTPTGTDVSLNFMPE 78

Db 351 LQLEPNHSGDYVLLSNMYASEQRWSDVQKIRKQMLRDGVKKV-----PG- 394

Qy 79 KRMHGIIFPKWVHE-----HPEITVCY-----QNDHIDLMKES 112

Db 395 ---HSLVEGVGNRVHFEFLMGDKSHQSDAIYAKUKEMTGRRLRSEGVYFQISNVYVDVEEE 451

Qy 113 REKYETYPKVV-----PEFAEITFLKN 135

Db 452 KENAVVHSEKIAIAFMLISTPERSPTTVKN 483

RESULT 5

US-11-056-355B-81255

;; Sequence 81255, Application US/11056355B

;; Publication No. US20060150283A1

;; GENERAL INFORMATION:

;; APPLICANT: Brover, Vyacheslav

;; APPLICANT: Alexandrov, Nikolai

;; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding

;; TITLE OF INVENTION: Polypeptides Encoded Thereby

;; FILE REFERENCE: 2750-1590PUS2

;; CURRENT APPLICATION NUMBER: US/11/056,355B

;; CURRENT FILING DATE: 2005-02-14

;; PRIOR APPLICATION NUMBER: 60/544,190


```
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 81255
; LENGTH: 1510
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: (1)..(1510)
; OTHER INFORMATION: Ceres Seq. ID no. 12659031
US-11-056-355B-81255

Query Match      8.8%; Score 78; DB 7; Length 1510;
Best Local Similarity 21.7%; Pred. No. 43;
Matches 33; Conservative 20; Mismatches 43; Indels 56; Gaps 6;

Qy 21 IYKNDHTIDYRIHSGMVAG--RWRDQEVNIVKLTGVYKVSWTEPTGTDVSLNFPNE 78
Db 406 LQLEPNHSGDYVLLSNMYASEQRWSDVQKIRKQMLRDGVKKV-----PG- 449
Qy 79 KRWHGIIFFPKWVHE-----HPEITVCY-----QNDHIDLAKES 112
Db 450 ---HSLVEGNVRVHEFLMGDKSHQSDAIYAKLKEMTGRLSRGVVPQISNVYVDVEEE 506
Qy 113 REKYETPKYVW-----PEFAEITFLKN 135
Db 507 KENAVVHSEKIAFAFMLISTPERSPTIVKN 538

RESULT 6
US-11-056-355B-81254
; Sequence 81254, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 81254
; LENGTH: 1523
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: (1)..(1523)
; OTHER INFORMATION: Ceres Seq. ID no. 12659030
US-11-056-355B-81254

Query Match      8.8%; Score 78; DB 7; Length 1523;
Best Local Similarity 21.7%; Pred. No. 44;
Matches 33; Conservative 20; Mismatches 43; Indels 56; Gaps 6;

Qy 21 IYKNDHTIDYRIHSGMVAG--RWRDQEVNIVKLTGVYKVSWTEPTGTDVSLNFPNE 78
Db 419 LQLEPNHSGDYVLLSNMYASEQRWSDVQKIRKQMLRDGVKKV-----PG- 462
Qy 79 KRWHGIIFFPKWVHE-----HPEITVCY-----QNDHIDLAKES 112
Db 463 ---HSLVEGNVRVHEFLMGDKSHQSDAIYAKLKEMTGRLSRGVVPQISNVYVDVEEE 519
Qy 113 REKYETPKYVW-----PEFAEITFLKN 135
Db 520 KENAVVHSEKIAFAFMLISTPERSPTIVKN 551

; RESULT 7
US-11-330-403-6834
```

```
; Sequence 6834, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 6834
; LENGTH: 354
; TYPE: prt
; ORGANISM: Neisseria meningitidis MC58
US-11-330-403-6834

Query Match      8.7%; Score 77; DB 7; Length 354;
Best Local Similarity 24.6%; Pred. No. 8.4;
Matches 33; Conservative 18; Mismatches 45; Indels 38; Gaps 6;

Qy 37 MVAGRWVRDQEVNIVKLT-----GVYKVSWTEPTGTDVSLNFPNEKRMHGIIFFP 88
Db 1 MKARFYDKGDIREDIPEPTVAPGTGVI-NVAVCGICGTDL-----HEFMGEPPIP 52
Qy 89 KWVHEHEPTVCYQNDHIDLAKESREKYEYTPKYVWPEFAEITFLKNEGVDNEV---IS 145
Db 53 PCGHPEP-----ISGESAPVTMGHEFGVGVVAVGEGVDDIKVGQHV 94
Qy 146 KAFYEGMTDIRAG 159
Db 95 VEPYI-IRDDVPTG 107

RESULT 8
US-11-330-403-4821
; Sequence 4821, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 4821
; LENGTH: 752
; TYPE: prt
; ORGANISM: Cytophaga hutchinsonii
US-11-330-403-4821

Query Match      8.7%; Score 76.5; DB 7; Length 752;
Best Local Similarity 21.0%; Pred. No. 25;
Matches 29; Conservative 29; Mismatches 53; Indels 27; Gaps 5;

Qy 1 MENFIGSHMIYTYENGWYEYIYKN-----DHTIDYRIHSGMVAGRWVRDQ 46
Db 188 LESFLGLHPHTDIPVQARELYIKNLLRVIGDVNYKVPYITIDDSQNQL-----DL 240
Qy 47 EVNIVKLTGVYKVSWTEPTGTDVSLNF-MPNEKRMHGII-----FPKRWHEHEPTVCY 101
Db 241 SCSVLRVSVPFH-VQYLNHTGVGATLTLSLHKKGLWGLVACHHYSKYLHYETKLAAL 299
Qy 102 QNDHIDLAKESREKYEY 119
Db 300 QGHFITSQIIEIQNEQY 317

RESULT 9
US-11-434-137-4210
; Sequence 4210, Application US/11434137
; Publication No. US20060210579A1
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
```

```

; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hervé
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,137
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 4210
; LENGTH: 1370
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-11-434-137-4210

```

	Query Match	8.7%	Score 76.5;	DB 7;	Length 1370;
	Best Local Similarity	20.3%;	Pred. No. 54;		
	Matches	28; Conservative	26; Mismatches	47; Indels	37; Gaps
QY	51 VKLTEGVYKYSWTEPTGTDVSL---NWPNEKR-----MHGIFF	87			
Dd	743 LKIIVDELUVKWGYEPEQIIVEMARENQTNGRRNSRORYKLDDGVKNLASDLGNLI--	800			
QY	88 PKVWEHPEITVCQNHDIDL--MKSEREKY-----ETYPKVVPFEAFITFLKEGV	138			
Dd	801 ---LKVEPTDQALQNERLFYYLYLQNGEDMYTGEALDIDNLSDQYDIHHIIQAIFKDSSI	857			
QY	139 DNEEVISKAPYEGMTDDI	156			
Dd	858 DNRLVLSSAKNRGKSDDV	875			

```

RESULT 10
US-11-434-184-4210
; Sequence 4210, Application US/11434184
; Publication No. US20060210580A1
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hevse
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,184
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 4210
; LENGTH: 1370
; TYPE: prt
; ORGANISM: Streptococcus agalactiae
US-11-434-184-4210

```

```

Query Match      8.7%; Score 76.5; DB 7; Length 1370;
Best Local Similarity 20.3%; Pred. No. 54;
Matches 28; Conservative 26; Mismatches 47; Indels 37; Gaps 5;

QY      51 VKLTGEVGVKVSWEPTGTGTVSL---NFMFNEKR-----MHGIFF 87
      :|: :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: |||
DB      743 LKIVDELVKVWGVEPQIUVEMARENOTTQGRNRSQRVKLLDDGVKNLASDLNGNI-- 800

QY      88 PKWVHHPELTVCYQNDHDL--MKESREXY-----ETYPKVVVEFABEITLKNQGV 138
      :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: |||
DB      801 ---LK EYPTDNQALQNERFLYYLQGRDNYMTGEALDIDLNSQYDIIHPQAFIKDSDI 857
      ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: |||

QY      139 DNEEVISKAPYEGMTDDI 156
      ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: |||
DB      858 DNRVLVSSAKNRCKSDDV 875
      ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: |||

RESULT 11
US-11-434-199-4210
; Sequence 4210, Application US/11434199
; Publication No. US20060210581A1

```

```

RESULT 11
US-11-434-199-4210
; Sequence 4210, Application US/11434199
; Publication No. US20060210581A1
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hervé
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,199
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: Seqwin99, version 1.02
; SEQ ID NO 4210
; LENGTH: 1370
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
; US-11-434-199-4210

```

[illegible]

RESULT 12
US-11-434-203-4210
; Sequence 4210, Application US/11434203
; Publication No. US20060210582A1
; GENERAL INFORMATION:
; APPLICANT: Telford, John

```

; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hervé
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,203
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 4210
; LENGTH: 1370
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-11-434-203-4210

Query Match      8.7%; Score 76.5; DB 7; Length 1370;
Best Local Similarity 20.3%; Pred. No. 54;
Matches 28; Conservative 26; Mismatches 47; Indels 37; Gaps 5;

QY 51 VKUTGYYKVSWTBPTGTDVSL---NFMNPKR-----MHGIIPF 87
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 743 LKIVDELVKVMGYEPEQIVEMARENOTTQGRNRSORYKLLDDGVKNLASDLNGNI-- 800

QY 88 PKWVHEHPETVCYONDHIDI--MKESREKY-----ETYPKYVVPFEAFITELKNEGV 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 801 ---LKEYPTDQNALQNERLFYLYLQNGRDMYTGEALDIDNLSQYDIDHIIPQAFIKDDSI 857

QY 139 DNEEVISKAPYEGMTDDI 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 858 DNRVLVSSAKNRGKSDDV 875

RESULT 13
US-10-449-902-47193
; Sequence 47193, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A020511-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47193
; LENGTH: 1077
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-47193

Query Match      8.5%; Score 75; DB 6; Length 1077;
Best Local Similarity 23.8%; Pred. No. 56;
Matches 40; Conservative 21; Mismatches 55; Indels 52; Gaps 7;

QY 12 TYENGWEYIYKNDHTIDYRIHSGWAGRWV--RDQEVNIVKLTEGVYKV----- 60

```

```

; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(422)
; OTHER INFORMATION: Ceres Seq. ID no. 12376550
US-11-056-355B-16491

Query Match      8.4%; Score 74.5; DB 7; Length 422;
Best Local Similarity 25.7%; Pred. No. 19;
Matches 26; Conservative 9; Mismatches 29; Indels 37; Gaps 5;

Qy 49 NIVKLTGEG-----VYKVSWTE-----PTGT-----DVSLNFMNPNEKRMHGIIFF 87
   :: ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: |||
Db 119 DVAGYTEGFPRLVYFEDSWTKVSNVIFLDAPVGTGFSYSREAGLNVSLSLTSGRQHHVFL 178
   :: ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: ||| :: |||

Qy 88 PKWVHEHPEITVCYONDHIDLKESREKY----ETYPKYVVP 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 RKWVAEHPEFA-----SNPLYIGGDSYSGYTVP 206

```

Search completed: October 12, 2006, 05:37:10
Job time : 17 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2006, 05:32:32 ; Search time 67 Seconds
(without alignments)
231.207 Million cell updates/sec

Title: US-10-824-581-4

Perfect score: 882

Sequence: 1 MENFIGSHMIYENGWEYE.....EVISKAPYEGMTDDIRAGRL 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_80.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	882	100.0	161	2 D69671	ferulate decarboxy
2	697	79.0	177	2 B86866	phenolic acid deca
3	419	47.5	174	2 B2100	decarboxylase VC22
4	91.5	10.4	1574	2 G70466	DNA-directed RNA p
5	82.5	9.4	312	2 T09443	zinc metalloprotei
6	82.5	9.4	312	2 B82276	hypothetical prote
7	82	9.3	1012	2 B30389	conserved hypothet
8	81	9.2	233	2 AC1176	carbonic anhydrase
9	81	9.2	511	2 T48777	6-HYDROXY-D-NICOTI
10	81	9.2	689	2 AF2122	prolyl endopeptida
11	80.5	9.1	273	2 A37245	probable Zn-depend
12	80	9.1	1035	2 T30287	exo-alpha-sialidas
13	80	9.1	1035	2 G98063	exo-alpha-sialidas
14	80	9.1	1402	2 A12816	DNA-directed RNA p
15	80	9.1	1416	2 B97595	hypothetical prote
16	79	9.0	523	2 T04825	hypothetical prote
17	78.5	8.9	464	2 AC2025	hypothetical prote
18	78.5	8.9	1000	2 T13636	probable minor str
19	78.5	8.9	1676	1 C5HU	complement C5 prec
20	78.5	8.9	1680	1 C5MU	complement C5 prec
21	78	8.8	233	2 AG1533	carbonic anhydrase
22	78	8.8	986	2 S49394	HsdR1 protein - My
23	78	8.8	1331	2 T04938	hypothetical prote
24	78	8.8	1495	2 A85240	hypothetical prote
25	78	8.8	1495	2 T10649	hypothetical prote
26	78	8.8	3135	2 A48584	transmission block
27	77.5	8.8	554	2 T25902	hypothetical prote
28	77	8.7	354	2 E81178	probable alcohol d
29	77	8.7	443	2 B90098	hypothetical prote

30	77	8.7	723	2	D42707	probable cation-tr
31	77	8.7	909	2	T00009	probable primase (
32	76	8.6	285	2	S77356	hypothetical prote
33	76	8.6	327	2	JT0584	deoxyribonuclease
34	76	8.6	354	2	H81925	probable alcohol d
35	76	8.6	854	1	WMBP22	gene 12 protein -
36	75.5	8.6	233	1	VHVUAV	nucleoprotein N -
37	75.5	8.6	236	2	G75324	hypothetical prote
38	75.5	8.6	477	2	B70363	conserved hypothet
39	75	8.5	190	2	AP2044	hypothetical prote
40	75	8.5	514	2	S54019	hypothetical prote
41	75	8.5	1452	2	T29107	polymorphic antige
42	74.5	8.4	405	2	E71179	hypothetical prote
43	74.5	8.4	470	2	T48003	hypothetical prote
44	74.5	8.4	534	2	T27588	hypothetical prote
45	74.5	8.4	686	2	C64428	hypothetical prote

ALIGNMENTS

RESULT 1

D69671

ferulate decarboxylase padC - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C/Accession: D69671

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: D69671

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-161 <KUN>

A:Cross-references: UNIPROT:O07006; UNIPARC:UPI0000060AB9; GB:Z99121; GB:AL009126; NID:9

A:Experimental source: strain 168

C:Genetics:

A:Gene: padC

C:Superfamily: phenolic acid decarboxylase

Query Match 100.0%; Score 882; DB 2; Length 161;

Best Local Similarity 100.0%; Pred. No. 3.2e-71;

Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MENFIGSHMIYENGWEYEIYIKNDHTIDYRHSAGVAGRWVDDQEVNIYKLTGVYKV 60

Db 1 MENFIGSHMIYENGWEYEIYIKNDHTIDYRHSAGVAGRWVDDQEVNIYKLTGVYKV 60

Qy 61 SWTEPTCTDYSLNFMNPKRMHGIIFPKVWVHEHPITVCYQNDHIDLMKESREKYETYP 120

Db 61 SWTEPTCTDYSLNFMNPKRMHGIIFPKVWVHEHPITVCYQNDHIDLMKESREKYETYP 120

Qy 121 KYVVPFAEITFLKNEGVDNEEVIKAPYEGMTDDIRAGRL 161

Db 121 KYVVPFAEITFLKNEGVDNEEVIKAPYEGMTDDIRAGRL 161

RESULT 2

B86866

phenolic acid decarboxylase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

RESULT 4

G70466
DNA-directed RNA polymerase (EC 2.7.7.6) beta' chain - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: G70466
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: G70466
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1574 <AOP>
A:Cross-references: UNIPROT:067763; UNIPARC:UPI0000056758; GB:AE000764; NID:g2984190; PI
A:Experimental source: strain VFS
C:Genetics:
A:Gene: rpoC
C:Superfamily: Escherichia coli DNA-directed RNA polymerase beta' chain
C:Keywords: nucleotidyltransferase; transcription

Query Match 10.4%; Score 91.5; DB 2; Length 1574;
Best Local Similarity 22.0%; Pred. No. 9.1;
Matches 39; Conservative 32; Mismatches 69; Indels 37; Gaps 6;

QY 4 FIGSHMITYENGWEYIYIKNDHTIDYRIHSGMVAGVRDQEVNIVKLTGKVYKVSMT 63
Db 134 YFESYLVITVEYPNEEEEEKFEKDEHTIP--LNDG-ISTKWVKLHVNEEFEE----- 182
QY 64 EPTGTDVSLNFPNPKRWHGI-----IFFPKVHHEHPEITVCYQNDHIDLKES 112
Db 183 -----KYAFTIDSEYHGMGAETLKEVLSKLDLDAVSRKLEIVKPYISGFEDLGKEI 235
QY 113 REK---YETPKYVPEP---ABITLKNQGVNNEEVISKAPYEGMTDDIRAGRL 161
Db 236 EQKYNLYQKLKVIADDPFAYGVKIGLEDHGLSLEQAIHRLNEELYLNVTGSI 292

RESULT 5

T09443
zinc metalloproteinase homolog - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T09443
R:Karolis, D.K.R.; Johnson, J.A.; Bailey, C.C.; Boedecker, E.C.; Kaper, J.B.; Reeves, P.
Proc. Natl. Acad. Sci. U.S.A. 95, 3134-3139, 1998
A:Title: A Vibrio cholerae pathogenicity island associated with epidemic and pandemic st
A:Reference number: Z16672; MUID:98169509; PMID:9501228
A:Accession: T09443
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-312 <KAP>
A:Cross-references: UNIPROT:068338; UNIPARC:UPI000000B9A07; EMBL:AF034434; NID:g3004923;
A:Experimental source: strain 16961

C;Genetics:
A;Note: part of the pathogenicity island (VPI); associated with epidemic and pandemic strains
C;Superfamily: Vibrio cholerae probable zinc metalloproteinase

Query Match 9.4%; Score 82.5; DB 2; Length 312;
Best Local Similarity 26.8%; Pred. No. 7.3;
Matches 45; Conservative 21; Mismatches 63; Indels 39; Gaps 10;

Qy 14 ENGW-----EYEIYIKNDHTIDYRIHSGMWAGWRDQEVNIV-----KLTEGVY 58
Db 131 KSGWNVTHEIKLVTHDNVTTYWEH---LKGPVPRGHEAGITWDSIPGAGSPGLTELVI 187

Qy 59 KV-----SWTPTPTGDVSLNFWPNE-----KRMHGIIFFPKWVHEHPBITVCYQNDDHD 107
Db 188 ALQKNPDGRKAAPTGNHGSKNLVIHFEGFHTLDRVVGTFTGKPYSQNPSF---YQAWhSD 244

Db 776 GYFITA KGA A I R S L E E N R V E I A Y B K S C T E N K O I R L C Y D Y L S N I Y L L E L R K T E S V 835
Qy 117 -----E T Y K Y V V P E A E I T F L K N E G V D N E E V I S K A P Y E 150
Db 836 I E T D K I V E I Y K L E V V E A K R H V K I N E K I E I K Y E I I H E 877

RESULT 8
AC1176
carbonic anhydrase homolog lmc0811 [imported] - Listeria monocytogenes (strain EGD-e)
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AC1176
R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, A.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mäkelä, C.; Schlüter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AC1176
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-233 <GLA>
A/Cross-references: UNIPROT:Q8Y8T3; UNIPARC:UPI000005535C; GB:NC_003210; PIDN:CAC98889.1
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmc0811

Query Match 9.2%; Score 81; DB 2; Length 233;
Best Local Similarity 23.7%; Pred. No. 6.8;
Matches 42; Conservative 24; Mismatches 69; Indels 42; Gaps 7;

Qy 11 Y Y E N G W E I Y I K N D H T T D Y R I H S G W A G R V R D E V N I V K L T E G V Y K S W T E ----- 64
Db 10 Y D E K T G P E M W G H I C S D F E I ---A H T G K -A Q S P V D I E Q A D V V K L K P S T M K F Y K E T D Y T I R 65
Qy 65 -----P T G T D V S L N F-----M P N E K R M H G I I F F P K W--V H E H P E I T V C 100
Db 66 R I E Q S V H V F P H D K E Q C L R E N G E Y Y P L V S P H A I P A E H L L D G Y I P I E W H F V H E K P D G T T L 125
Qy 101 Y Q N D H I D L M K E R E K Y E T Y P K V V P E F A-----E I T F L K N E G V D N E E V I S K A P Y E G 151
Db 126 V M S A W M D I N T N N I E F K N L P T Y F P E V A D F E T E R E I T L D V N E F M P E R V F--Y T Y Q G 180

RESULT 9
T48777
6-HYDROXY-D-NICOTINE OXIDASE related protein [imported] - Neurospora crassa
N/Alternate names: protein 13E11.250
C/Species: Neurospora crassa
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C/Accession: T48777
R/Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, A.; Reference number: Z24541
A/Accession: T48777
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-511 <SCH>
A/Cross-references: UNIPROT:Q9P621; UNIPARC:UPI000006C774; EMBL:AL353820; GSPDB:GN00112
A/Experimental source: cosmid contig 13E11, strain 74
C/Genetics:
A/Gene: NCSP:13E11.250
A/Map position: 2
A/Introns: 32/3
C/Superfamily: poppy reticuline oxidase

Query Match 9.2%; Score 81; DB 2; Length 511;
Best Local Similarity 28.4%; Pred. No. 18;
Matches 25; Conservative 19; Mismatches 32; Indels 12; Gaps 4;

QY 69 DVSINFPNKRMGHGIIFPPKWHVHEHPEITVCYO--ND--HIDLMKBSREKYEYTPKYVV 124
 Db 264 NLTVNGMPSDNKAHSYM-----VLTHMDMLGGYAVISDLFHADLPNTADSAHETVPEVFV 318
 QY 125 PEFAEITFLKNEGVDNEEISKA---PY 149
 Db 319 PFESLPITTKSIDNVSTISRAIDVPY 346

RESULT 10
 AF2122
 prolyl endopeptidase [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AF2122
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AF2122
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-689 <KUR>
 A:Cross-references: UNIPROT:O8YU28; UNIPARC:UPI00000CE482; GB:BA000019; PIDN:BA074232.1
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: ali2533
 C:Superfamily: prolyl oligopeptidase

Query Match 9.2%; Score 81; DB 2; Length 689;
 Best Local Similarity 21.9%; Pred. No. 27;
 Matches 30; Conservative 18; Mismatches 45; Indels 44; Gaps 6;

QY 14 ENGWEYEYIYKNDHTIDYRIHSGMVAG--RVVRDQEVNIVKLTEGVYKVSMTPTGTDVS 71
 Db 5 EKSINYPUSHKIDHVDY--HGTLVADPYRLDPP-----SEKTRV--- 44

QY 72 LNFMNPKRMHGIIFPPKWHVHEHPEITVCYQNDHI--DLMKESREKYEYTPKYVVP--EF 127
 Db 45 -----WIEAQNOITPAYLGEVSIREKIQRNLKLDVYKYGIPPKEG 86

QY 128 AEITFLKNEGVDNEEVI 144
 Db 87 ENYFYFKNDGLQNSVL 103

RESULT 11
 A97245
 probable Zn-dependent hydrolase from metallo-beta-lactamase superfamily [imported] - Clostridium acetobutylicum
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C:Accession: A97245
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: A97245
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-273 <KUR>
 A:Cross-references: UNIPROT:Q97FD6; UNIPARC:UPI00000CA33C; GB:AE001437; PIDN:AAK80748.1
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC2804

Query Match 9.1%; Score 80.5; DB 2; Length 273;
 Best Local Similarity 23.4%; Pred. No. 9.2;
 Matches 37; Conservative 25; Mismatches 35; Indels 61; Gaps 11;

QY 17 WEYEYI-----KNDHTI-DYRIHSGMVAGRVVRDQEVNIVKLTEGVYKVSMTPE 65
 Db 93 WSPVVIHELEMPYVVGKDYDLADPTVDEGMYA-----KISOTFP 133
 QY 66 -TGTDSLNF--MPNKRMHGIIFFP--KVVHE--HPE--ITVCYQNDHIDLMKES----- 112
 Db 134 HTSIDISSNVSLPDS---SIFYPDWKIHTPGTGDGHISLFRKDXHLVIVGDAFCTT 190
 QY 113 -----REKYETPKYVVPPEAE-----ITFLKN 135
 Db 191 QBSLLSVITHHEEISGPPKYLTTDWAKAKNSIICLKN 228

RESULT 12
 T30287
 exo-alpha-sialidase (EC 3.2.1.18) - Streptococcus pneumoniae
 N:Alternate names: neuraminidase
 C:Species: Streptococcus pneumoniae
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
 C:Accession: T30287
 R:Camara, M.; Boulnois, G.J.; Andrew, P.W.; Mitchell, T.J.
 Infect. Immun. 62, 3688-3695, 1994
 A:Title: A neuraminidase from Streptococcus pneumoniae has the features of a surface protein
 A:Reference number: Z20807; MUID:94341870; PMID:8063384
 A:Accession: T30287
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1035 <CAM>
 A:Cross-references: UNIPROT:Q59959; UNIPARC:UPI000012FD46; EMBL:X72967; NID:9587550; PID
 C:Genetics:
 A:Gene: nana
 C:Keywords: glycosidase; hydrolase

Query Match 9.1%; Score 80; DB 2; Length 1035;
 Best Local Similarity 23.7%; Pred. No. 56;
 Matches 36; Conservative 30; Mismatches 64; Indels 22; Gaps 7;

QY 7 SHMITYTENG--WEYEYIYKNDHTID--YRIHSGMVAGRVVRDQEVNIVKLTEGVYKVSMT 63
 Db 604 SRIYSDHDKTWHAGEAVNDNRQVDQKIHSSNMNRRAQNTSTVVQLNNGDVKLFMR 663

QY 64 EPTGTVSLNFMNPKRMHGIIFPPKWHVHEHPEITVCY--QNDHIDLMKESRE----- 114
 Db 664 GLTG-DLQV-----ATSKDGGVTWEKDKRYQPKVQVYVQMSAIHTMHEGKEYIILSNAG 717

QY 115 --KYETPKYV--VPEFAEITFLKNEGVDNEE 142
 Db 718 GPKRENGMVHLARVEENGELTWLKHNPQKGE 749

RESULT 13
 G98063
 exo-alpha-sialidase (EC 3.2.1.18) precursor (neuraminidase A) [imported] - Streptococcus pneumoniae
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C:Accession: G98063
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Bargett, S.; DeHoff, B.S.; E., R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; N., P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.; J. Bacteriol. 183, 5709-5717, 2001
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: G98063
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1035 <KUR>
 A:Cross-references: UNIPROT:Q59959; UNIPARC:UPI000012FD46; GB:AE007317; PIDN:AA00340.1;
 C:Genetics:
 A:Gene: nana
 C:Keywords: glycosidase; hydrolase

Query Match 9.1%; Score 80; DB 2; Length 1035;


```
Best Local Similarity 23.7%; Pred. No. 56;
Matches 36; Conservative 30; Mismatches 64; Indels 22; Gaps 7;

QY 7 SHMITYENG--WEVEIYKNDHTD--YRIHSGMVAGRWDRDQEVNIVKLTEGVYKVSMT 63
DB 604 SRIIYSDHGTWTHAGEAVNDRQVDGQKIHSSTNNRRQAQNTSTVVLNNGDVKLFMR 663

QY 64 EPTGTDVSLNFMNPKRMHGIIFPPKWHPEITVCY-QNDHIDLMKESRE----- 114
DB 664 GLTG-DLQV-----ATSKDGGVTWEKDKRPQVKDVVQMSAHTMHEGKEYIILSNAG 717

QY 115 --KYETPKYV--VPEFAEITFLKNEGVDNEE 142
DB 718 GPKRENGMVHLARVEENGELTWLKHNPQKGE 749

RESULT 14
AI2816
DNA-directed RNA polymerase beta' chain [imported] - Agrobacterium tumefaciens (strain C
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AI2816
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AI2816
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1402 <KUR>
A:Cross-references: UNIPROT:Q8UE09; UNIPARC:UPI0000164673; GB:AE008688; PIDN:AAL42951.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: rpoC
A:Map position: circular chromosome
C:Superfamily: Escherichia coli DNA-directed RNA polymerase beta' chain

Query Match 9.1%; Score 80; DB 2; Length 1402;
Best Local Similarity 22.0%; Pred. No. 82;
Matches 40; Conservative 23; Mismatches 69; Indels 50; Gaps 7;

QY 17 WEVEIYKNDHTIDYRI-----HSGMVAGRW---VRDQEVNIVKLTEGVYKVSMTPTGTD 69
DB 34 WSYG-EIKKPETINRYTFKPERDGLFCARIFGPINKDYECCLCGKYKRMKYKGIICEKCGVE 92

QY 70 VSLNFMNPKRMH-----GIIFFPKWVHEHPEI 97
DB 93 VTLRVRERMGHIELAAPVAHIWFLKSLPSRISTLLDMLTKDVERVLYFENYIVTEPGL 152

QY 98 TVCYQNDHIDLMKESREKYETPKYVPEFAEITFLKNEGVDN-BEVISKAPYEGMTDDI 156
DB 153 TSLKQNQLL-----SEEEY---MIAVDFEGEDQFTAMIGAEIYEMLSMNLKXIAGDL 203

QY 157 RA 158
DB 204 RA 205

RESULT 15
B97595
Hypothetical protein AGR_C3568 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: B97595
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
```

```
A:Accession: B97595
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1416 <KUR>
A:Cross-references: UNIPROT:Q8UE09; UNIPARC:UPI000000D1D23; GB:AE007869; PIDN:AAK87715.1;
C:Genetics:
A:Gene: AGR_C3568
A:Map position: circular chromosome
C:Superfamily: Escherichia coli DNA-directed RNA polymerase beta' chain

Query Match 9.1%; Score 80; DB 2; Length 1416;
Best Local Similarity 22.0%; Pred. No. 83;
Matches 40; Conservative 23; Mismatches 69; Indels 50; Gaps 7;

QY 17 WEVEIYKNDHTIDYRI-----HSGMVAGRW---VRDQEVNIVKLTEGVYKVSMTPTGTD 69
DB 48 WSYG-EIKKPETINRYTFKPERDGLFCARIFGPINKDYECCLCGKYKRMKYKGIICEKCGVE 106

QY 70 VSLNFMNPKRMH-----GIIFFPKWVHEHPEI 97
DB 107 VTLRVRERMGHIELAAPVAHIWFLKSLPSRISTLLDMLTKDVERVLYFENYIVTEPGL 166

QY 98 TVCYQNDHIDLMKESREKYETPKYVPEFAEITFLKNEGVDN-BEVISKAPYEGMTDDI 156
DB 167 TSLKQNQLL-----SEEEY---MIAVDFEGEDQFTAMIGAEIYEMLSMNLKXIAGDL 217

QY 157 RA 158
DB 218 RA 219

Search completed: October 12, 2006, 05:33:47
Job time : 68 secs
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2006, 05:32:05 ; Search time 144 Seconds
(without alignments)
1034.219 Million cell updates/sec

Title: US-10-824-581-4

Perfect score: 882

Sequence: 1 MENFIGSHMYTVYENGWEYE.....EVISKAPYEGWTDIDIRAGL 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	882	100.0	161	1	PADC_BACSU	O07006 bacillus su
2	863	97.8	161	2	O8KNX7	O8KNX7 bacillus sp
3	800	90.7	166	2	O65FC9	O65FC9 bacillus li
4	777	88.1	161	2	O45381	O45381 bacillus pu
5	769	87.2	161	2	O9EXR7	O9EXR7 bacillus pu
6	697	79.0	177	2	O9CEB3	O9CEB3 lactococcus
7	695	78.8	178	2	O9F3X2	O9F3X2 pediococcus
8	689	78.1	178	2	O38UX6	O38UX6 lactobacill
9	689	78.1	178	2	O38UX6	O38UX6 lactobacill
10	675	76.5	176	2	O3Y2T7	O3Y2T7 enterococcu
11	670	76.0	178	2	O88R77	O88R77 lactobacill
12	666.5	75.6	174	2	P94900	P94900 lactobacill
13	493	55.9	108	2	O9KH18	O9KH18 lactobacill
14	493	55.9	109	2	O9KH17	O9KH17 lactobacill
15	489	55.4	177	2	O743A0	O743A0 mycobacteri
16	488	55.3	107	2	O9KHJ0	O9KHJ0 lactobacill
17	474	53.7	108	2	O9KHJ0	O9KHJ0 lactobacill
18	467	52.9	107	2	O9KH16	O9KH16 lactobacill
19	464	52.6	109	2	O9KHJ1	O9KHJ1 lactobacill
20	461	52.3	169	2	O6DB32	O6DB32 erwiniac car
21	419	47.5	174	1	PADC_VIBCH	O9AGX3 vibrio chol
22	329	37.3	95	2	O9KHJ2	O9KHJ2 lactobacill
23	189	21.4	39	2	O9R4C1	O9R4C1 bacillus pu
24	172.5	19.6	38	2	O9R4W3	O9R4W3 pseudomonas
25	138	15.6	176	2	O4IH86	O4IH86 gibberella
26	115	13.0	172	2	O2U7L8	O2U7L8 aspergillus
27	112.5	12.8	175	2	O4P8S8	O4P8S8 ustilago ma
28	94	10.7	681	2	O43KC0	O43KC0 chlorobium
29	92	10.4	866	2	O6FSX2	O6FSX2 candida gla
30	92	10.4	1504	2	O6B961	O6B961 candida gla
31	91.5	10.4	1574	1	RPOC_AQUAE	O67763 aquifex aeo

32	86.5	9.8	312	2	O9AGX3	VIBCH
33	86.5	9.8	1588	2	O50SL7	ENTHI
34	85	9.6	379	2	O8THI9	METAC
35	85	9.6	540	2	O6BIU2	DEBHA
36	84.5	9.6	462	2	O32LT4	BRARE
37	84.5	9.6	527	2	O6FOT4	MESFL
38	84.5	9.6	1044	2	O5KYP9	GEOKA
39	84	9.5	379	2	O46DI3	METBA
40	84	9.5	400	2	O4ER00	LISMO
41	83.5	9.5	196	2	O537O2	9ERYO
42	83.5	9.5	350	2	O5T089	HUMAN
43	83.5	9.5	350	2	O8WM30	HUMAN
44	83.5	9.5	366	2	O5OR13	ENTHI
45	83.5	9.5	497	2	O9H852	HUMAN

ALIGNMENTS

RESULT 1
PADC_BACSU STANDARD; PRT; 161 AA.
AC O07006;
DT 10-OCT-2002, integrated into UniProtKB/Swiss-Prot.
DT 01-JUL-1997, sequence version 1.
DE 07-MAR-2006, entry version 37.
DE Phenolic acid decarboxylase [EC 4.1.1.-] (PAD).
GN Name=padC; Synonym=pad; OrderedLocusNames=BSU34400;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND CHARACTERIZATION.
RC STRAIN=168;
RX MEDLINE=98207851; PubMed=9546183;
RA Cavin J.-F., Dartois V., Divies C.;
RT "Gene cloning, transcriptional analysis, purification, and
characterization of phenolic acid decarboxylase from Bacillus
subtilis";
RT Appl. Environ. Microbiol. 64:1466-1471 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC DENIZOT F.;
RA DENIZOT F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
Borries R., Boursier L., Brang A., Braun M., Brignell S.C., Bron S.,
Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
Choi S.-K., Codani J.-J., Connerton I.F., Cummings N.J., Daniel R.A.,
Denizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emmerson P.T.,
Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
Ghim S.-Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
Guiseppi G., Guy B.-J., Haga K., Haiech J., Harwood C.R., Henaut A.,
Hilbert H., Holsappel S., Hosono S., Hulio M.-F., Itaya M.,
Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Mael C.,
Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S.,
Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
Shin B.-S., Soldo B., Sorokin A., Taccioni E., Takagi T., Takahashi H.,
Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P.,

Togononi A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yaumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumstein E., Yoshikawa H., Danchin A.,
 RA "the complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*,"
 RL Nature 390:249-256(1997).
 CC -!- FUNCTION: Catalyzes the decarboxylation of phenolic acids such as ferulic, p-coumaric and caffeic acids.
 CC -!- BIOPHYSICO-CHEMICAL PROPERTIES:
 CC pH dependence:
 CC Optimum pH is 5.0;
 CC Temperature dependence:
 CC Optimum temperature is 40-45 degrees Celsius;
 CC -!- SUBUNIT: Homodimer (Probable).
 CC -!- INDUCTION: By ferulic, p-coumaric and caffeic acids. Cells extracts from caffeic acid-induced cells exhibited lower activity on the three acids, which indicates that caffeic acid could be a less efficient inducer.
 CC -!- SIMILARITY: Belongs to the padC family.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

 DR EMBL; AF017117; AAC46254.1; -; Genomic DNA.
 DR EMBL; Z94043; CAB08020.1; -; Genomic DNA.
 DR EMBL; Z99121; CAB15445.1; -; Genomic DNA.
 DR PIR; D69671; D69671.
 DR GenomeReviews; AL009126 GR; BSU34400.
 DR Subtilist; BG12433; padC.
 DR BioCyc; BSUB1423:BSU3437-MONOMER; --
 DR InterPro; IPR008729; PA decarbox.
 DR Pfam; PF05870; PA decarbox; 1.
 DR PIRSF; PIRSF011561; PAD; 1.
 DR ProDom; PD022010; PA decarbox; 1.
 DR Complete proteome; Decarboxylase; Lyase.
 KW CHAIN 1 161 Phenolic acid decarboxylase.
 FT /FTID=PRO_0000108125.
 FT SEQUENCE 161 AA; 19077 MW; BAF73F679D0FC313 CRC64;
 Query Match 100.0%; Score 882; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred. No. 1.8e-70;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKV 60
 DB 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKV 60
 QY 61 SWTEPTGTDVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLMKESREKYETYP 120
 DB 61 SWTEPTGTDVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLMKESREKYETYP 120
 QY 121 KYVVPFEAITEFLKNEGVDNEEVISKAPYEGMTDDIRAGRL 161
 DB 121 KYVVPFEAITEFLKNEGVDNEEVISKAPYEGMTDDIRAGRL 161
 RESULT 2
 ID Q8KNX7_9BACI PRELIMINARY; PRT; 161 AA.
 AC Q8KNX7;
 DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2002, sequence version 1.
 DE 07-FEB-2006, entry version 11.
 DE Phenolic acid decarboxylase.
 GN Name=padA;
 OS *Bacillus* sp. BP-7.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 ON NCBI_TaxID=126733;
 RX (1)
 RP NUCLEOTIDE SEQUENCE.
 MEDLINE=22947447; PubMed=12819959; DOI=10.1007/s00253-003-1371-y;
 RA "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*,"
 RL Nature 390:249-256(1997).
 CC -!- FUNCTION: Catalyzes the decarboxylation of phenolic acids such as ferulic, p-coumaric and caffeic acids.
 CC -!- BIOPHYSICO-CHEMICAL PROPERTIES:
 CC pH dependence:
 CC Optimum pH is 5.0;
 CC Temperature dependence:
 CC Optimum temperature is 40-45 degrees Celsius;
 CC -!- SUBUNIT: Homodimer (Probable).
 CC -!- INDUCTION: By ferulic, p-coumaric and caffeic acids. Cells extracts from caffeic acid-induced cells exhibited lower activity on the three acids, which indicates that caffeic acid could be a less efficient inducer.
 CC -!- SIMILARITY: Belongs to the padC family.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

 DR EMBL; AF017117; AAC46254.1; -; Genomic DNA.
 DR EMBL; Z94043; CAB08020.1; -; Genomic DNA.
 DR EMBL; Z99121; CAB15445.1; -; Genomic DNA.
 DR PIR; D69671; D69671.
 DR GenomeReviews; AL009126 GR; BSU34400.
 DR Subtilist; BG12433; padC.
 DR BioCyc; BSUB1423:BSU3437-MONOMER; --
 DR InterPro; IPR008729; PA decarbox.
 DR Pfam; PF05870; PA decarbox; 1.
 DR PIRSF; PIRSF011561; PAD; 1.
 DR ProDom; PD022010; PA decarbox; 1.
 DR Complete proteome; Decarboxylase; Lyase.
 KW CHAIN 1 161 Phenolic acid decarboxylase.
 FT /FTID=PRO_0000108125.
 FT SEQUENCE 161 AA; 19077 MW; BAF73F679D0FC313 CRC64;
 Query Match 100.0%; Score 882; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred. No. 1.8e-70;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKV 60
 DB 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKV 60
 QY 61 SWTEPTGTDVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLMKESREKYETYP 120
 DB 61 SWTEPTGTDVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLMKESREKYETYP 120
 QY 121 KYVVPFEAITEFLKNEGVDNEEVISKAPYEGMTDDIRAGRL 161
 DB 121 KYVVPFEAITEFLKNEGVDNEEVISKAPYEGMTDDIRAGRL 161
 RESULT 2
 ID Q8KNX7_9BACI PRELIMINARY; PRT; 161 AA.
 AC Q8KNX7;
 DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2002, sequence version 1.
 DE 07-FEB-2006, entry version 11.
 DE Phenolic acid decarboxylase.
 GN Name=padA;
 OS *Bacillus* sp. BP-7.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 ON NCBI_TaxID=126733;
 RX (1)
 RP NUCLEOTIDE SEQUENCE.
 MEDLINE=22947447; PubMed=12819959; DOI=10.1007/s00253-003-1371-y;

Prim N., Pastor F.I.J., Diaz P.;
 RA "Biochemical studies on cloned *Bacillus* sp. BP-7 phenolic acid decarboxylase PadA,"
 RL Appl. Microbiol. Biotechnol. 63:51-56(2003).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

 DR EMBL; AJ492219; CAD37333.1; -; Genomic DNA.
 DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
 DR InterPro; IPR008729; PA decarbox.
 DR Pfam; PF05870; PA decarbox; 1.
 DR PIRSF; PIRSF011561; PAD; 1.
 DR ProDom; PD022010; PA decarbox; 1.
 DR SEQUENCE 161 AA; 19134 MW; 4B870F7BED06B2AE CRC64;
 Query Match 97.8%; Score 863; DB 2; Length 161;
 Best Local Similarity 97.5%; Pred. No. 8.8e-69;
 Matches 157; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKV 60
 DB 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKV 60
 QY 61 SWTEPTGTDVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLMKESREKYETYP 120
 DB 61 SWTEPTGTDVSLNFMNPKRMHGIIFPPKVVHEHPEITVCYQNDHIDLMKESREKYETYP 120
 QY 121 KYVVPFEAITEFLKNEGVDNEEVISKAPYEGMTDDIRAGRL 161
 DB 121 KYVVPFEAITEFLKNEGVDNEEVISKAPYEGMTDDIRAGRL 161
 RESULT 3
 ID Q65FC9_BACLD PRELIMINARY; PRT; 166 AA.
 AC Q65FC9; Q620U4;
 DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
 DT 25-OCT-2004, sequence version 1.
 DT 07-FEB-2006, entry version 12.
 DE PadC (Phenolic acid decarboxylase).
 GN Name=padC; OrderedLocNames=BL02117, BLI03407;
 OS *Bacillus licheniformis* (strain DSM 13 / ATCC 14580).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 ON NCBI_TaxID=279010;
 RX (1)
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15383718; DOI=10.1159/000079829;
 RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H., Ehrenreich P., Baumer S., Henne A., Liesegang H., Merkl R., Ehrenreich A., Gottschalk G.;
 RA "The complete genome sequence of *Bacillus licheniformis* DSM13, an organism with great industrial potential,"
 RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 RX (2)
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;
 RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J., Tang M., Lopez de Leon A., Xiang H., Gusti V., Clausen I.G., Olsen P.B., Rasmussen M.D., Andersen J.T., Joergensen P.B., Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D., Berka R.M.;
 RA "Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species.";
 RL Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

 DR EMBL; AE017333; AAU42235.1; -; Genomic DNA.
 DR EMBL; CP000002; AAU24866.1; -; Genomic DNA.
 DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
 DR InterPro; IPR008729; PA decarbox.

```

DR Pfam; PF05870; PA decarbox; 1.
DR PIRSF; PIRSF011561; PAD; 1.
DR ProDom; PD022010; PA decarbox; 1.
KW Complete proteome.
SQ SEQUENCE 166 AA; 19522 MW; 3F966922733353ED CRC64;

Query Match      90.7%; Score 800; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 3.8e-63;
Matches 140; Conservative 14; Mismatches 6; Indels 0; Gaps 0;

QY 1 MENFIGSHMITYENGWEYIYKNDHTIDYRIHSGMWAGRWDRDQEVNIVKLTGKVYK 60
DB ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
5 VKFVGSMMITYENGWEYIYKNDHTIDYRIHSGMWAGRWDRDQEVNIVKLTGKVYK 64
QY 61 SWTEPTGTVSLNPMPEKRMHGIIFFPKWVHEPITVCYQNDHIDLMKESREKYETP 120
DB ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
65 SWTEPTGTVSLNPMPEKRMHGIIFFPKWVHEPITVCYQNDHIDLMKESREKYETP 124
QY 121 KYVPEFAEITFLKNEGVDNEEVISKAPYEGMTDDIRAGR 160
DB ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
125 KYVPEFADITFIENAGDINEDLSKAPYPGWTDIDIRAGK 164

RESULT 4
Q45361 BACPU
ID Q45361 BACPU PRELIMINARY; PRT; 161 AA.
AC Q45361;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1996, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Ferulate decarboxylase.
GN Name=fdc;
OS Bacillus pumilus (Bacillus mesentericus);
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1408;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PS213;
RX MEDLINE=96086033; PubMed=8534115;
DOI=10.1128/AEM.67.3.1063-1069.2001;
RA Zago A., Degross G., Bruschi C.V.;
RT "Cloning, sequencing and expression in E. coli of the Bacillus pumilus
RT gene for ferulic acid decarboxylase.";
RL Appl. Environ. Microbiol. 61:4484-4486 (1995).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; X84815; CAA59273.1; -; Genomic DNA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR008729; PA decarbox.
DR Pfam; PF05870; PA decarbox; 1.
DR PIRSF; PIRSF011561; PAD; 1.
DR ProDom; PD022010; PA decarbox; 1.
SQ SEQUENCE 161 AA; 19082 MW; 074F9DCF1990B07F CRC64;

Query Match      88.1%; Score 777; DB 2; Length 161;
Best Local Similarity 85.0%; Pred. No. 4.1e-61;
Matches 136; Conservative 16; Mismatches 8; Indels 0; Gaps 0;

QY 1 MENFIGSHMITYENGWEYIYKNDHTIDYRIHSGMWAGRWDRDQEVNIVKLTGKVYK 60
DB ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1 MDQFVGLHMIYTYENGWEYIYKNDHTIDYRIHSGMWAGRWDRDQEVNIVKLTGKVYK 60
QY 61 SWTEPTGTVSLNPMPEKRMHGIIFFPKWVHEPITVCYQNDHIDLMKESREKYETP 120
DB ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
121 KYVPEFAEITFLKNEGVDNEEVISKAPYEGMTDDIRAGR 160
DB ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
125 KYVPEFADITFIENAGDINEDLSKAPYPGWTDIDIRAGK 160

RESULT 5
Q45361 BACPU
ID Q45361 BACPU PRELIMINARY; PRT; 161 AA.
AC Q45361;
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2001, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Phenolic acid decarboxylase (PAD).
GN Name=pad;
OS Bacillus pumilus (Bacillus mesentericus);
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1408;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15884;
RX MEDLINE=21141815; PubMed=11229892;
DOI=10.1128/AEM.67.3.1063-1069.2001;
RA Barthelmebs L., Davies C., Cavin J.F.;
RT "Expression in Escherichia coli of native and chimeric phenolic acid
RT decarboxylases with modified enzymatic activities and method for
RT screening recombinant E. coli strains expressing these enzymes.";
RL Appl. Environ. Microbiol. 67:1063-1069 (2001).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; AJ278683; CAC18719.1; -; Genomic DNA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR008729; PA decarbox.
DR Pfam; PF05870; PA decarbox; 1.
DR PIRSF; PIRSF011561; PAD; 1.
DR ProDom; PD022010; PA decarbox; 1.
SQ SEQUENCE 161 AA; 19032 MW; 6566415CA65AD19D CRC64;

Query Match      87.2%; Score 769; DB 2; Length 161;
Best Local Similarity 84.4%; Pred. No. 2.1e-60;
Matches 135; Conservative 16; Mismatches 9; Indels 0; Gaps 0;

QY 1 MENFIGSHMITYENGWEYIYKNDHTIDYRIHSGMWAGRWDRDQEVNIVKLTGKVYK 60
DB ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1 MDQFVGLHMIYTYENGWEYIYKNDHTIDYRIHSGMWAGRWDRDQEVNIVKLTGKVYK 60
QY 61 SWTEPTGTVSLNPMPEKRMHGIIFFPKWVHEPITVCYQNDHIDLMKESREKYETP 120
DB ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
121 KYVPEFAEITFLKNEGVDNEEVISKAPYEGMTDDIRAGR 160
DB ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
125 KYVPEFADITFIENAGDINEDLSKAPYPGWTDIDIRAGK 160

RESULT 6
Q45361 BACPU
ID Q45361 BACPU PRELIMINARY; PRT; 177 AA.
AC Q45361;
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Phenolic acid decarboxylase.
GN Name=fdc; OrderedLocustNames=LL1930;
OS Lactococcus lactis subsp. lactis (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=LL1403;
RX MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr-GR-1697R;
RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis LL1403.";
RL Genome Res. 11:731-753 (2001).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms

```

```
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AE006423; AAK06028.1; -; Genomic_DNA.
DR PIR; B86866; B86866.
DR Biocyc; LLAC1360.L193734-MONOMER; -.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR008729; PA decarbox.
DR Pfam; PF05870; PA decarbox; 1.
DR PIRSF; PIRSF011561; PAD; 1.
DR ProDom; PD022010; PA decarbox; 1.
DR Complete proteome.
SQ SEQUENCE 177 AA; 20892 MW; 11C6AC123FB6B683 CRC64;

Query Match 79.0%; Score 697; DB 2; Length 177;
Best Local Similarity 71.9%; Pred. No. 6.1e-54;
Matches 115; Conservative 31; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MENFIGSHMITYENGWEYIYKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKV 60
Db 7 LEDFVGTHTFYTDNGWEYELVKNHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKI 66
Qy 61 SWTEPTGTVSLNFMNPKRMHGIIFPKWVHEHPEITVCYQNDHIDLMKESREKYETYP 120
Db 67 TWTEPTGTVSLNFMNPKRMHGIIFPKWVHEHPEITVCYQNDHIDLMKESREKYETYP 126
Qy 121 KYVPEFAEITFLKNEGVDNEEVISKAPYEGMTDDIRAGR 160
Db 127 KYLVPEFAKITVAEAGKNDODVVAQAQPYKEMTNDIRNGK 166

RESULT 7
Q9F3X2 PEDPE
ID Q9F3X2_PEDPE PRELIMINARY; PRT; 178 AA.
AC Q9F3X2;
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2001, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Phenolic acid decarboxylase.
GN Name=padA;
OS Pedicoccus pentosaceus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae; Pediococcus.
OX NCBI_TaxID=1255;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=20528324; PubMed=11073918;
RX DOI=10.1128/JB.182.23.6724-6731.2000;
RA Barthelmebs L., Lecomte B., Divies C., Gavin J.F.;
RT "Inducible metabolism of phenolic acids in Pediococcus pentosaceus is encoded by an autoregulated operon which involves a new class of negative transcriptional regulator.";
RL J. Bacteriol. 182:6724-6731(2000).
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AJ276891; CAC16794.1; -; Genomic DNA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR008729; PA decarbox.
DR Pfam; PF05870; PA decarbox; 1.
DR PIRSF; PIRSF011561; PAD; 1.
DR ProDom; PD022010; PA decarbox; 1.
DR Complete proteome.
SQ SEQUENCE 178 AA; 20873 MW; 130A5D55B441065F CRC64;

Query Match 78.8%; Score 695; DB 2; Length 178;
Best Local Similarity 73.8%; Pred. No. 9.3e-54;
Matches 118; Conservative 27; Mismatches 15; Indels 0; Gaps 0;

Qy 1 MENFIGSHMITYENGWEYIYKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKV 60
Db 8 LDDFLGTHFYTDNGWEYEWAKNDHTVDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKI 67
Qy 61 SWTEPTGTVSLNFMNPKRMHGIIFPKWVHEHPEITVCYQNDHIDLMKESREKYETYP 120
Db 127 KYLVPEFAKITVAEAGKNDODVVAQAQPYKEMTNDIRNGK 166

us-10-824-581-4.rup
```

```
Db 68 ANTEPTGTVSLNFMNPKRMHGIIFPKWVHEHPEITVCYQNDHIDLMKESREKYETYP 127
Qy 121 KYVPEFAEITFLKNEGVDNEEVISKAPYEGMTDDIRAGR 160
Db 128 KLVVPEFATITMGDAGQNDNEVIAEAPYEGMTDDIRAGK 167

RESULT 8
Q93MP3 LACSK
ID Q93MP3_LACSK PRELIMINARY; PRT; 178 AA.
AC Q93MP3;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE P-coumaric acid decarboxylase PdcA.
GN Name=pdcA;
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OX NCBI_TaxID=1599;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=21822499; PubMed=11832506;
RX Duzet A.-M., Chaillou S., Hissler L., Stentz R.,
RA Champomier-Verges M.-C., Albert C.-A., Zagorec M.;
RT "Physical and genetic map of the Lactobacillus sakei 23K chromosome.";
RL Microbiology 148:421-431(2002).
RN NUCLEOTIDE SEQUENCE.
RP Duzet A.-M., Chaillou S., Zagorec M.;
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF400064; AAK85433.1; -; Genomic DNA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR008729; PA decarbox.
DR Pfam; PF05870; PA decarbox; 1.
DR PIRSF; PIRSF011561; PAD; 1.
DR ProDom; PD022010; PA decarbox; 1.
DR Complete proteome.
SQ SEQUENCE 178 AA; 21040 MW; 07708F630910470D CRC64;

Query Match 78.1%; Score 689; DB 2; Length 178;
Best Local Similarity 73.8%; Pred. No. 3.2e-53;
Matches 118; Conservative 24; Mismatches 18; Indels 0; Gaps 0;

Qy 1 MENFIGSHMITYENGWEYIYKNDHTIDYRIHSGMVAGRWVRDQEVNIVKLTEGVYKV 60
Db 8 LEDFLGTHTFYTDNGWYEWYAKNDHTVDYRIHSGMVAGRWVRDQEVNIVKLTDGVFKI 67
Qy 61 SWTEPTGTVSLNFMNPKRMHGIIFPKWVHEHPEITVCYQNDHIDLMKESREKYETYP 120
Db 68 TWTEPTGTVSLNFMNPKRMHGIIFPKWVHEHPEITVCYQNDHIDLMKESREKYETYP 127

us-10-824-581-4.rup
```

```
RESULT 9
Q38UX6 LACSS
ID Q38UX6_LACSS PRELIMINARY; PRT; 178 AA.
AC Q38UX6;
DT 22-NOV-2005, integrated into UniProtKB/TrEMBL.
DT 22-NOV-2005, sequence version 1.
DT 07-MAR-2006, entry version 4.
DE P-coumaric acid decarboxylase.
GN Name=pdcA; OrderedLocNames=LSA1701;
OS Lactobacillus sakei subsp. sakei (strain 23K).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OX NCBI_TaxID=314315;

us-10-824-581-4.rup
```

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP PubMed:16273110; DOI=10.1038/nbt1160;
 RA Chailion S., Champomier-Verges M.-C., Cornet M., Crutz-Le Coq A.-M.,
 RA Duzet A.-M., Martin V., Beaufils S., Darbon-Rongere E., Bossy R.,
 RA Loux V., Zagorec M.,
 RT "The complete genome sequence of the meat-borne lactic acid bacterium
 RT Lactobacillus sakei 23K".
 RL Nat. Biotechnol. 23:1527-1533 (2005).
 CC
 CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL: CR936503; CAI56008.1; -: Genomic DNA.
 DR GenomeReviews: CR936503 GR; LSA1701.
 DR GO: GO:0016831; F:carboxy-lyase activity; IEA.
 KW Complete proteome.
 SQ SEQUENCE 178 AA; 21040 MW; 07708F630910470D CRC64;

 Query Match 78.1%; Score 689; DB 2; Length 178;
 Best Local Similarity 73.8%; Pred. No. 3.2e-53;
 Matches 118; Conservative 24; Mismatches 18; Indels 0; Gaps 0;

 Qy 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDQEVNIIVKLTEGVYKV 60
 Db 8 LEFLGTHFIYTDNGWYEWYAKNDHTVDYRIHSGMVAGRWVRDQEVNIIVKLTEGVYKV 67

 Qy 61 SWTEPTGTVSLNMPNEKRMHGIIFPPKWHHEHPEITVYQNDHIDLMKESRKYETYP 120
 Db 68 TWTEPTGTVSLNMPNEKRMHGIIFPPKWHHEHPEITVYQNDHIDLMKESRKYETYP 127

 Qy 121 KYVPEAEITFLKNEGVNDEEISKAPYEGMTDIRAGR 160
 Db 128 KLLVPEFATITMGDAGNEDVISQAPYEGMTDIRAGR 167

RESULT 10
 Q3Y2T7.ENTFC
 ID Q3Y2T7.ENTFC PRELIMINARY; PRT; 176 AA.
 AC Q3Y2T7;
 DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 11-OCT-2005, sequence version 1.
 DT 07-FEB-2006, entry version 3.
 DE Phenolic acid decarboxylase.
 GN ORFNames=EfaeDRAFT_2126;
 OS Enterococcus faecium DO.
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 ON NCBI_TaxID=333849;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=DO;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Israni S., Pittluck S., Richardson P.,
 RT "Sequencing of the draft genome and assembly of Enterococcus faecium
 RT DO".
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RP
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=DO;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome assembly of Enterococcus faecium DO";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RP
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=DO;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Israni S., Pittluck S., Richardson P.;
 RT Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 WC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.
 CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL: AAA03000004; EAM10835.1; -: Genomic DNA.
 DR GO: GO:0016831; F:carboxy-lyase activity; IEA.
 DR InterPro: IPR008729; PA decarbox.
 DR Pfam: PF05870; PA decarbox; 1.
 DR PIRSF: PIRSF011561; PAD; 1.
 DR PRODOM: PD022010; PA decarbox; 1.
 SQ SEQUENCE 176 AA; 20800 MW; 31787220E5A8E83 CRC64;

 Query Match 76.5%; Score 675; DB 2; Length 176;
 Best Local Similarity 70.6%; Pred. No. 5.5e-52;
 Matches 113; Conservative 30; Mismatches 17; Indels 0; Gaps 0;

 Qy 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDQEVNIIVKLTEGVYKV 60
 Db 7 LEFLGTHFIYTDNGWYEWYAKNDHTVDYRIHSGMVAGRWVRDQEVNIIVKLTEGVYKV 66

 Qy 61 SWTEPTGTVSLNMPNEKRMHGIIFPPKWHHEHPEITVYQNDHIDLMKESRKYETYP 120
 Db 67 TWTEPTGTVSLNMPNEKRMHGIIFPPKWHHEHPEITVYQNDHIDLMKESRKYETYP 126

 Qy 121 KYVPEAEITFLKNEGVNDEEISKAPYEGMTDIRAGR 160
 Db 127 KLLVPEFATITMGDAGNEDVISQAPYEGMTDIRAGR 166

 RESULT 11
 Q88RY7.LACPL
 ID Q88RY7.LACPL PRELIMINARY; PRT; 178 AA.
 AC Q88RY7;
 DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2003, sequence version 1.
 DT 07-FEB-2006, entry version 11.
 DE P-coumaric acid decarboxylase.
 GN Name=cdc; OrderedLocusNames=lp_3665;
 OS Lactobacillus plantarum.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 ON NCBI_TaxID=1590;
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=NCIMB 8826 / WCFS1;
 RC MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;
 RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
 RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
 RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
 RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
 RA De Vos W.M., Siezen R.J.;
 RT "Complete genome sequence of Lactobacillus plantarum WCFS1";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995 (2003).
 CC
 CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL: AL935262; CAD65735.1; -: Genomic DNA.
 DR BioCyc: LPLA220668:LP_3665-MONOMER; -;
 DR GO: GO:0016831; F:carboxy-lyase activity; IEA.
 DR InterPro: IPR008729; PA decarbox.
 DR Pfam: PF05870; PA decarbox; 1.
 DR PIRSF: PIRSF011561; PAD; 1.
 DR PRODOM: PD022010; PA decarbox; 1.
 KW Complete proteome.
 SQ SEQUENCE 178 AA; 21010 MW; C8ED70783E6F7496 CRC64;

 Query Match 76.0%; Score 670; DB 2; Length 178;
 Best Local Similarity 71.2%; Pred. No. 1.6e-51;
 Matches 114; Conservative 27; Mismatches 19; Indels 0; Gaps 0;

 Qy 1 MENFIGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDQEVNIIVKLTEGVYKV 60

```
Db 8 LDDFLGTHFIYTDNGWEYEWAKNDHTVDYRIHSGMVAGRWVTDQKADIVNLTGEGYKI 67
Qy 61 SWTEPTGTDVSLNFMNFKRMHGIFFPKWVHEHEPEITVCYQNDHIDLMKESREKYETYP 120
Db 68 SWTEPTGTDVSLNFMNFKRMHGIFFPKWVHEHEPEITVCYQNDHIDLMKESREKYATYP 127
Qy 121 KVVPEFAEITLKNQGVNDNEEIVSKAPYEGMTDDIRAGR 160
Db 128 KLVVPEFANITYMGDQGNEDVISEAPYKEMPNDIRNGK 167

RESULT 12
P94900 LACPL PRELIMINARY; PRT; 174 AA.
AC P94900;
DT 01-MAY-1997, integrated into UniProtKB/TrEMBL.
DT 01-MAY-1997, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE P-coumaric acid decarboxylase.
GN Name=cdc;
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97288084; PubMed=9143125;
RA Cavin J.F., Barthelmebs L., Davies C.;
RT "Molecular characterization of an inducible p-coumaric acid
RT decarboxylase from Lactobacillus plantarum: Gene cloning,
RT transcriptional analysis, overexpression in Escherichia coli,
RT purification, and characterization.";
RL Appl. Environ. Microbiol. 63:1939-1944 (1997).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; U63827; AAC45282.1; -; Genomic DNA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR008729; PA_decarbox.
DR Pfam; PF05870; PA_decarbox; 1.
DR PIRSF; PIRSF011561; PAD; 1.
DR ProDom; PD022010; PA_decarbox; 1.
DR SEQUENCE 174 AA; 20406 MW; 7ABEFC7374A4B42C4 CRC64;

Query Match 75.6%; Score 666.5; DB 2; Length 174;
Best Local Similarity 71.9%; Pred. No. 3.1e-51;
Matches 115; Conservative 26; Mismatches 18; Indels 1; Gaps 1;

Qy 1 MENFIGSHMIYTYNGWEYEWAKNDHTVDYRIHSGMVAGRWVTDQKADIVNLTGEGYKI 67
Db 8 LDDFLGTHFIYTDNGWEYEWAKNDHTVDYRIHSGMVAGRWVTDQKADIVNLTGEGYKI 67
Qy 61 SWTEPTGTDVSLNFMNFKRMHGIFFPKWVHEHEPEITVCYQNDHIDLMKESREKYETYP 120
Db 68 SWTEPTGTDVSLNFMNFKRMHGIFFPKWVHEHEPEITVCYQNDHIDLMKESREKYATYP 127
Qy 121 KVVPEFAEITLKNQGVNDNEEIVSKAPYEGMTDDIRAGR 160
Db 128 KLVVPEFANITYMGDQGNEDVISEAPYKEMPNDIRNGK 166

RESULT 13
Q9KH18 LACFE PRELIMINARY; PRT; 108 AA.
AC Q9KH18;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Phenolic acid decarboxylase (Fragment).
OS Lactobacillus fermentum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
```

```
OC Lactobacillus.
OX NCBI_TaxID=1613;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA van Beek S., Priest F.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF257162; AAF82764.1; -; Genomic DNA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR008729; PA_decarbox.
DR Pfam; PF05870; PA_decarbox; 1.
DR ProDom; PD022010; PA_decarbox; 1.
FT NON TER 108
FT NON TER 108
SQ SEQUENCE 108 AA; 12719 MW; 75679E90AEB869B7 CRC64;

Query Match 55.9%; Score 493; DB 2; Length 108;
Best Local Similarity 78.7%; Pred. No. 5.2e-36;
Matches 85; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

Qy 18 EYEIYKNDHTIDYRIHSGMVAGRWVTDQKADIVNLTGEGYKI SWTEPTGTDVSLNFMN 77
Db 1 EYEWAKNDHTVDYRIHSGMVAGRWVTDQKADIVNLTGEGYKI SWTEPTGTDVSLNFMN 60
Qy 78 EKRMHGIIFFPKWVHEHEPEITVCYQNDHIDLMKESREKYETYPKYVVP 125
Db 61 EKHLGHTIFFPKWVHEHEPEITVCYQNDHIDLMKESREKYATYPKLVVP 108

RESULT 14
Q9KH17 LACPL PRELIMINARY; PRT; 109 AA.
AC Q9KH17;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Phenolic acid decarboxylase (Fragment).
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA van Beek S., Priest F.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF257163; AAF82765.1; -; Genomic DNA.
DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
DR InterPro; IPR008729; PA_decarbox.
DR Pfam; PF05870; PA_decarbox; 1.
DR ProDom; PD022010; PA_decarbox; 1.
FT NON TER 109
FT NON TER 109
SQ SEQUENCE 109 AA; 12833 MW; 38D13D1EB52D8E29 CRC64;

Query Match 55.9%; Score 493; DB 2; Length 109;
Best Local Similarity 78.7%; Pred. No. 5.3e-36;
Matches 85; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

Qy 18 EYEIYKNDHTIDYRIHSGMVAGRWVTDQKADIVNLTGEGYKI SWTEPTGTDVSLNFMN 77
Db 2 EYEWAKNDHTVDYRIHSGMVAGRWVTDQKADIVNLTGEGYKI SWTEPTGTDVSLNFMN 61
Qy 78 EKRMHGIIFFPKWVHEHEPEITVCYQNDHIDLMKESREKYETYPKYVVP 125
```

Db 62 EKLLHGTTFFPKWVEEHPEITVYQNEHIDLMEQSKYATYPKLVVP 109

RESULT 15

Q743A0 MYCPA
ID Q743A0 MYCPA PRELIMINARY; PRT; 177 AA.
AC Q743A0
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Hypothetical protein.
GN OrderedLocNames=MAP0685; ORFNames=MAP_0685;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=k10;
RA Li L., Bamanatine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.

CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----

DR EMBL; AE016959; AAS03002.1; -; Genomic DNA.

DR GO; GO:0016831; F:carboxy-lyase activity; IEA.

DR InterPro; IPR008729; PA decarbox.

DR Pfam; PF05870; PA decarbox; 1.

DR ProDom; PD022010; PA decarbox; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 177 AA; 20326 MW; 36B84719009F648C CRC64;

Query Match 55.4%; Score 489; DB 2; Length 177;
Best Local Similarity 57.6%; Pred. No. 2.1e-35;
Matches 83; Conservative 28; Mismatches 33; Indels 0; Gaps 0;

Qy 5 IGSHMIYTYENGWEYIYIKNDHTIDYRIHSGMVAGRWVRDOEVNIVKLTGVYKVSWTE 64

Db 22 VGHRFPIYTYANGWQYEMVYKNATTIDYRIHSGHVGGRWVGQEVNVLQVDDDDSYKISWTE 81

Qy 65 PTGTDVSLNFMNPKRMHGIIFFPKWVHEHPEITVYQNDHIDLMEQSKYATYPKYV 124

Db 82 PTGTCVAVNVLPSKERIHGVIFFFQWIRQHGERTVCFQNEHLDENRAYDRGPTYPIYEV 141

Qy 125 PEFABITFLKNEGVNDNEEVISKAP 148

Db 142 PEFAYITLFEYVGTDETVIDTAP 165

Search completed: October 12, 2006, 05:34:40
Job time : 147 secs